

<https://helda.helsinki.fi>

---

## A unified framework for analysis of individual-based models in ecology and beyond

Cornell, Stephen J.

2019-10-17

---

Cornell , S J , Suprunenko , Y F , Finkelshtein , D , Somervuo , P & Ovaskainen , O 2019 , ' A unified framework for analysis of individual-based models in ecology and beyond ' , Nature Communications , vol. 10 , 4716 . <https://doi.org/10.1038/s41467-019-12172-y>

---

<http://hdl.handle.net/10138/307001>

<https://doi.org/10.1038/s41467-019-12172-y>

---

cc\_by

publishedVersion

---

*Downloaded from Helda, University of Helsinki institutional repository.*

*This is an electronic reprint of the original article.*

*This reprint may differ from the original in pagination and typographic detail.*

*Please cite the original version.*

ARTICLE

<https://doi.org/10.1038/s41467-019-12172-y>

OPEN

# A unified framework for analysis of individual-based models in ecology and beyond

Stephen J. Cornell<sup>1,6\*</sup>, Yevhen F. Suprunenko<sup>1,2,6</sup>, Dmitri Finkelshtein<sup>3</sup>, Panu Somervuo<sup>4</sup> & Otso Ovaskainen<sup>4,5</sup>

Individual-based models, ‘IBMs’, describe naturally the dynamics of interacting organisms or social or financial agents. They are considered too complex for mathematical analysis, but computer simulations of them cannot give the general insights required. Here, we resolve this problem with a general mathematical framework for IBMs containing interactions of an unlimited level of complexity, and derive equations that reliably approximate the effects of space and stochasticity. We provide software, specified in an accessible and intuitive graphical way, so any researcher can obtain analytical and simulation results for any particular IBM without algebraic manipulation. We illustrate the framework with examples from movement ecology, conservation biology, and evolutionary ecology. This framework will provide unprecedented insights into a hitherto intractable panoply of complex models across many scientific fields.

<sup>1</sup>Institute of Integrative Biology, University of Liverpool, Liverpool L69 7ZB, UK. <sup>2</sup>Department of Plant Sciences, University of Cambridge, Downing Street, Cambridge CB2 3EA, UK. <sup>3</sup>Department of Mathematics, Swansea University, Fabian Way, Swansea SA1 8EN, UK. <sup>4</sup>Organismal and Evolutionary Biology Research Programme, University of Helsinki, Helsinki P.O. Box 65, FI-00014, Finland. <sup>5</sup>Centre for Biodiversity Dynamics, Department of Biology, Norwegian University of Science and Technology, N-7491 Trondheim, Norway. <sup>6</sup>These authors contributed equally: Stephen J. Cornell, Yevhen F. Suprunenko.

\*email: [stephen.cornell@liverpool.ac.uk](mailto:stephen.cornell@liverpool.ac.uk)

There are many systems in biology<sup>1</sup>, the social sciences<sup>2</sup> and other disciplines comprising populations of individuals that move and interact with each other, but where the properties of focal interest arise at the whole-population level. In ecology and evolution<sup>1,3,4</sup>, for instance, large-scale patterns, such as coexistence, biodiversity, and the evolution of novel types emerge from individual-level processes, such as births, deaths and inter-agent interactions. Historically, however, many fundamental ideas in ecology were developed using differential equation models that treat populations as continuous numbers.

The mathematical transparency of these models means that we understand their predictions completely and can propose general principles, e.g., regarding the number of resources needed by competing species<sup>5</sup> or the processes that destabilise host–parasite dynamics<sup>6</sup>. However, it has long been recognised that these simple models are inadequate for describing many biological phenomena<sup>7</sup>. For example, the conditions for competitors to coexist, and for predator–prey cycles to occur, are altered when space and individual discreteness are considered<sup>8</sup>. Space and stochasticity can similarly play an important role in the dynamics of chemical species<sup>9</sup>. Individual-based (or agent-) based models<sup>10</sup> faithfully capture the discrete and spatial nature of population dynamics, but these are usually studied by computer simulation<sup>1</sup>, which only tells us about a limited set of parameter values and not the general model behaviour. They therefore do not yield the sort of general understanding that would be given by mathematical analysis, which encapsulates behaviour across all possible parameter combinations<sup>8</sup>.

To use individual-based models to develop general principles, of the sort derived from classic differential equation models, requires a method for analysing them mathematically. Individual-based models can be formulated by spatiotemporal point processes<sup>11</sup>, where individuals (or, more generally, entity types, such as juveniles or infected individuals, see Fig. 1a) are created, destroyed and move at rates that can depend on the positions of other individuals in the system, see Fig. 1b. In principle, the dynamics of such systems are described exactly by equations for the time evolution of spatial moments, representing mean density of individuals (first-order moment), spatial covariance (second-order moment) and so on (Fig. 1c). However, in practice there are two obstacles to using these to predict model behaviour. First, they need to be laboriously derived separately for each model, and for each order of moment<sup>12</sup>. Second, the moment equations form an unclosed hierarchy, with the dynamics of each moment depending on the higher order moments, so if the equations for all orders of moment were known it would still not be possible to solve them—even numerically<sup>13</sup>. ‘Moment closure’ is a widely used approximation scheme that closes the hierarchy by an ad-hoc assumption relating moments of different order<sup>12</sup>, but this is an uncontrolled approximation that is not guaranteed to perform well in any particular limit<sup>8,14</sup>. A more reliable alternative based on a perturbation expansion has been proposed, which gives asymptotically exact results when agents interact over large enough scales, but the algebraic burden for this method remains prohibitive because it requires arduous derivations for each particular model<sup>11,13</sup> (Fig. 1d).

Here, we overcome these difficulties by formulating a unified theoretical framework for a wide class of systems, which allows us to derive general analytical results. The results for any particular model within this general class can be obtained directly without further analysis. We make analytical results available to the non-specialist by providing software, which generates mathematical expressions describing a model that the user specifies in an accessible and intuitive way. We first describe our framework and the mathematical results leading to it, and then give three applications that illustrate how our method allows us to answer

questions that are not addressable by a simulation approach alone.

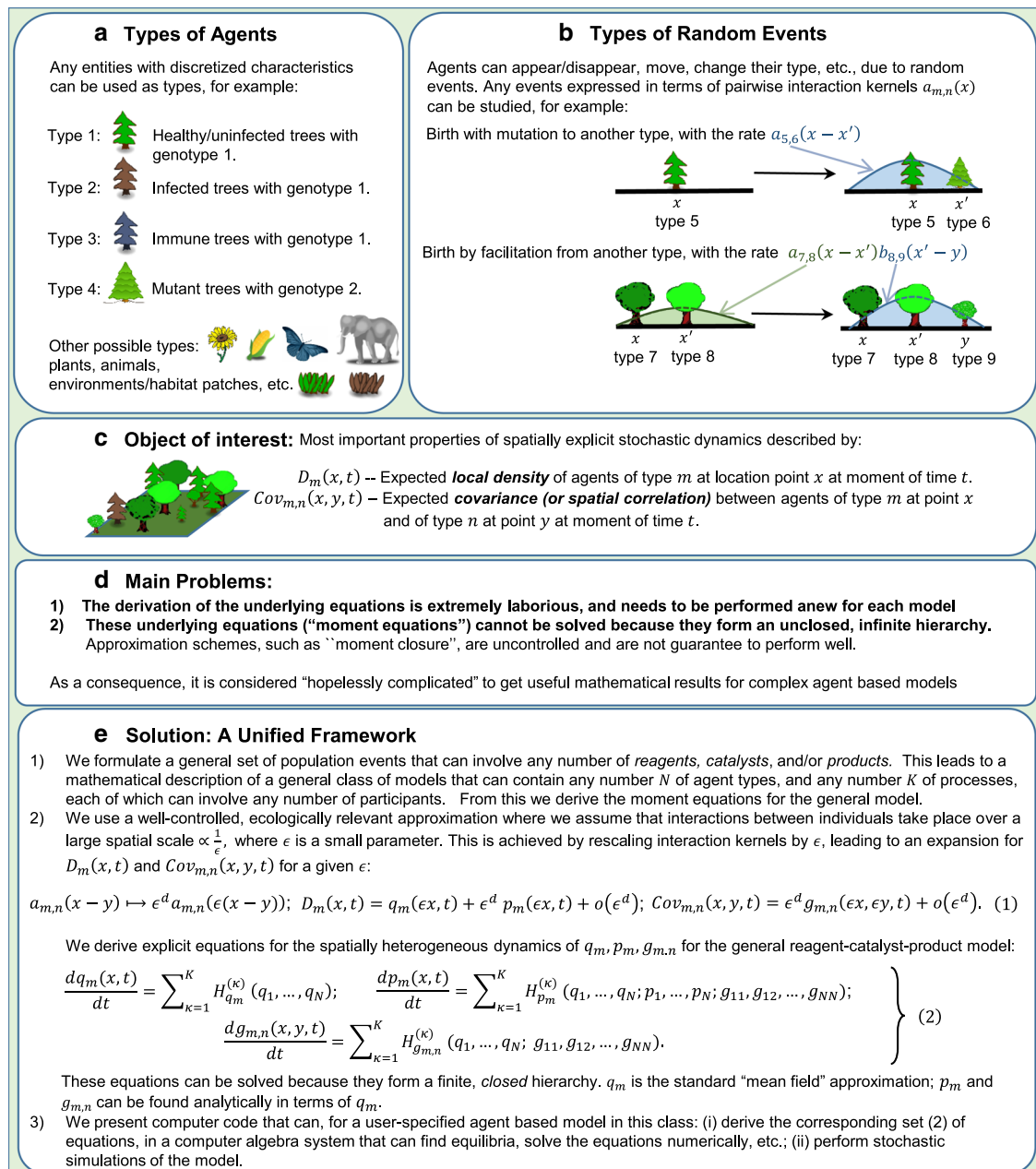
## Results

**The framework.** We begin by classifying the participants in demographic processes into three types of individuals (borrowing terminology from chemistry): (i) reactants (that are destroyed by the process); (ii) products (that are created by the process) and (iii) catalysts (that are unaffected by the process, but whose presence affects the rate at which it occurs). For example: a death event has a single reactant and no products or catalysts; dioecious birth has two catalysts (the mother and the father) and one product (the offspring) and movement can be represented by a reactant at the initial position and a product at the final position. This representation can describe processes with an arbitrarily high degree of complexity: the number of reactants, catalysts and products that can participate in any event is unlimited, as is the number of entity types within the system. We could, for instance, model a population where individuals can consume a food item, thereby increasing their energy level, provided they are in the shelter of a tree and that there is another individual helping to capture the prey. In this case, there would be two catalysts (tree and helper), two reactants (original individual and food item) and one product (an individual with a higher energy level than the original), and different entity types would be used to represent trees, food items and individuals of the focal species with different energy levels. This general classification allows us to derive our first main result: an exact expression for the moment equations to all orders for the general model containing processes with arbitrary sets of reactants, products and catalysts and interactions between them (see the Methods section Eq. (3)).

Then, using this exact expression for the general model, we apply a perturbation scheme<sup>13,15</sup> to derive a robust approximation for the dynamics of this general class of model. If the interactions between individuals take place over very large spatial scales, the demographic rates depend only on globally averaged population densities. In that case, the moment equations reduce to ‘mean-field’ equations—ordinary differential equations of the type used in classical population dynamics. If the interactions are more local, then stochastic and spatial fluctuations cause the model to deviate from the mean-field behaviour. We assume that spatial interactions between individuals depend on their separation  $x$  (in  $d$ -dimensional space) according to interaction ‘kernels’ of the form  $e^{df(ex)}$ , where  $1/\epsilon$  is the typical length scale of interactions, and prove mathematically (Supplementary Note 1) that the mean densities and spatial covariance (including autocovariance) satisfy an expansion

$$\begin{aligned}\text{density} &= q + \epsilon^d p + o(\epsilon^d), \\ \text{spatial covariance} &= \epsilon^d g(ex) + o(\epsilon^d),\end{aligned}$$

where  $q$  is the density computed by the mean-field model,  $p$  is the correction to the mean-field density due to spatial stochastic fluctuations and  $g$  is the dominant contribution to the spatial covariance that describes the degree to which individuals are aggregated or segregated in space ( $o(\epsilon^d)$  denotes a term that, when divided by  $\epsilon^d$ , vanishes when  $\epsilon \rightarrow 0$ ). For simplicity, in the above expression we have assumed translational invariance so that the density does not depend on position and the spatial covariance depends only on spatial separation, but our underlying mathematical framework does not have this limitation, see Fig. 1e and Eq. (1) therein. Depending on the scientific question being asked, it may be enough to study the mean-field behaviour described by  $q$  alone, whereas  $p$  or  $g$  will need to be calculated if we are interested in the nature and consequences of spatial patterns that emerge from local interactions. This perturbation

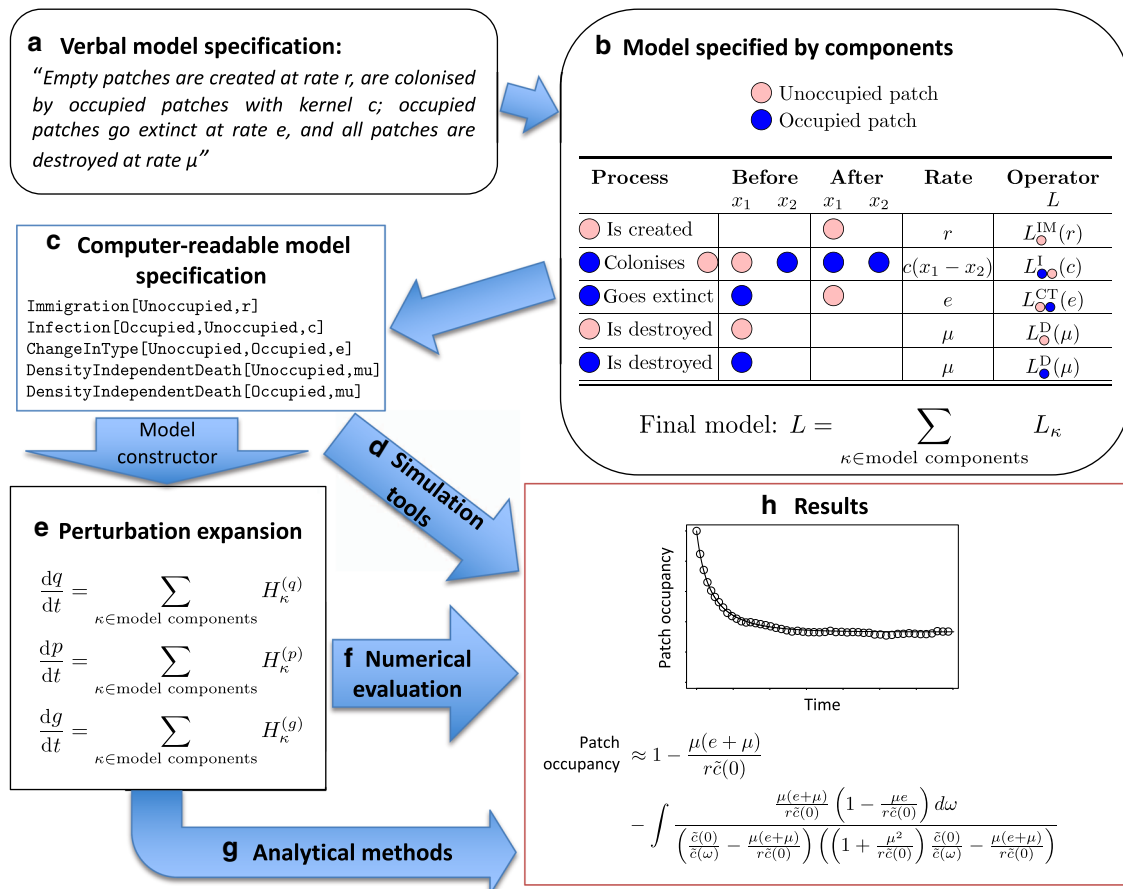


**Fig. 1** Agent-based models in ecology and evolution: the problem, and our solution. Agent-based models consist of entities of different types (panel **a**) that can interact in space in many ways (panel **b**). At any time, the state of the model can be quantified by spatial moments, such as density or spatial correlations (panel **c**). Agent-based models are generally considered too complicated to analyse mathematically (panel **d**). We address this problem by defining a general class of models and processes, for which we provide computer code that yields analytical solutions and runs stochastic simulations (panel **e**)

scheme has been applied to some specific ecological models before, and has been found to perform as well as or better than alternative moment closure approaches<sup>13,15</sup>. The framework described here allows us to apply this technique to a much wider class of models.

Our second main result is an expression for differential equations for  $q$ ,  $p$  and  $g$  for the general reactant–catalyst–product process (Fig. 1e and Eq. (2) therein, and Eqs. (4), (5), (6) in the Methods section). This allows us to construct directly the perturbation expansion for any model containing any number of processes in our general reactant–catalyst–product class because, even though the interactions between individuals may be nonlinear, the differential equations for  $q$ ,  $p$  and  $g$  contain

sums of independent contributions from each process. Our results make mathematical analysis available for systems that would be prohibitively complex using other approaches, and, while Eqs. (4), (5), (6) may appear daunting, we have automated this process by writing Mathematica code that computes analytical expressions for general reactant–catalyst–product models (see Supplementary Note 2). To complement and verify the analytical results, we have also written computer code in C that simulates a very broad class of reactant–catalyst–product models (see Supplementary Note 2). This provides a unified framework for analysis and simulation (illustrated in Fig. 2), where the user wanting to study a particular model has only to convert the verbal model description into a graphical model



**Fig. 2** Steps in our framework for constructing models and calculating their predictions. These are illustrated using our dynamic landscape metapopulation case study. The user wanting to study a particular model has only to convert the verbal model description (a) into a graphical model description (b), formulated in terms of a list of possible events and the rates at which they occur. An unambiguous mathematical statement of the model as a spatiotemporal point process is obtained from the operators  $L$  for these components<sup>11</sup>; the superscripts denote the corresponding class of reactant-catalyst-product process, namely immigration (IM), infection (I), change of type (CT), density-independent death (D). The colour denotes the type of individual (blue: occupied patch; pink: unoccupied patch). A computer-readable statement of the components (c) can be used to generate code for simulating the model (d), or to derive the equations describing the perturbation expansion using computer symbolic algebra (e). The differential equations for the mean-field  $q$ , correction  $p$  and spatial covariance  $g$  in the perturbation expansion comprise a sum of terms contributed by each of the model components. These expressions can then be evaluated numerically (f), or analysed further by human or machine (g). The predictions from simulations and mathematical methods can then be compared (h) and used in many ways to gain insights into the focal problem (see Figs. 3, 4 and 5)

description, which can easily be translated to the syntax understood by the computer code. For simplicity, our software assumes a translational invariant initial condition, because there are many ways in which the underlying geometry could be spatially heterogeneous (see the Methods section for more discussion of this point). Below we provide a step-by-step description of the application of the framework to derive the underlying equations for a dynamic landscape patch occupancy metapopulation model. We then use three case studies drawn from different areas of ecology and evolution to illustrate how the framework of Fig. 2 allows researchers to obtain deep analytical insights into important problems.

**Using the framework.** The use of our framework follows the workflow illustrated in Fig. 2. We demonstrate this with an example for which the perturbation expansion has previously been derived via other methods: a dynamic landscape patch occupancy metapopulation<sup>15</sup>. For simplicity, we will here consider the case without landscape correlations (the limit  $\nu \rightarrow 0$ ,  $\alpha\nu = \text{constant}$  in ref. 15).

We begin by deriving the equations for the sub-model that describes the dynamics of the habitat patches themselves (without

considering whether they are occupied), which is explicitly solvable. We assume that individual patches are created, at random, at rate per unit area  $r$ , and destroyed (at random) at rate  $\mu$ . Within our framework, this can be represented by a spatiotemporal point process with a single-entity type (habitat patches, which we denote as species type ‘1’). The formal mathematical representation of this model is the generator

$$L = L_1^{\text{IM}}(r) + L_1^{\text{D}}(\mu),$$

where the generators  $L^{\text{IM}}$  and  $L^{\text{D}}$  have technical definitions in the Supplementary Note 1, section 1.3, and can be understood intuitively as representing an immigration process (patch creation) with intensity  $r$ , and a density-independent death process (patch destruction) with rate  $\mu$ . The subscript ‘1’ states the species to which the process applies (in this case, habitat patches are the only ‘species’). These are two examples of the reactant-catalyst-product processes that, for convenience, we have defined in Supplementary Note 1 section 1.3, but note that we have derivations for the general reactant-catalyst-product process (Supplementary Note 1 section 1.4)—we could specify the model explicitly in terms of such general processes, but this would be somewhat more long-winded than using the above shorthand.



We will now show the steps for deriving the equations for this model using our Model Constructor (Supplementary Note 2 section 2.3). First, load the Mathematica packages containing the pre-defined list of processes and for deriving the equations:

```
get["SSPPLibraryOfProcesses"]
get["SSPPAnalyticalExpressions"]
```

and define the Mathematica vector describing the processes in the model:

```
processes={Immigration[1,r,1],DensityIndependentDeath[1,mu,1]}
```

The two functions `Immigration[]` and `DensityIndependentDeath[]` are pre-defined representations of the corresponding processes (see Supplementary Table 3 for the full list of pre-defined processes, and Supplementary Note 2 sections 2.1.3 and 2.3.4 for details on how to define specific reactant–catalyst–product processes). The first argument ‘1’ is the species label, the second argument ( $r$  or  $\mu$ ) is the corresponding rate and the third process is a (redundant) overall prefactor for the rates, included for consistency with other processes.

The quantity  $H_q$  (the right-hand side for the differential equation for the mean-field patch density) is given by the Mathematica command:

```
HQfALL[{q,p,g},processes,1]
>r - mu q[1]
```

where  $\{q,p,g\}$  specifies the names of the variables Mathematica will use to denote the mean-field density, stochastic correction and correlation functions, `processes` is the vector defining the model given above, and ‘1’ is the species label. To obtain the quantities  $H_p$ , which requires an integration in Fourier space, and  $H_g$ , we need to define the spatial dimension, and specify the name of the integration variable and the name for the frequency which is the argument of  $g$ :

```
HPfALL[{q,p,g},processes,1,k,2]
>0
HGfALL[{q,p,g},processes,1,1,k]
>0
```

For `HPfALL`, the argument ‘1’ refers to the species label (of which there is only one in this simple model), and the argument ‘2’ refers to the spatial dimension (which needs to be specified because `HPfALL` will in general include an integral over space). In `HGfALL`, both arguments ‘1’ denote species labels, of which there need to be two because this quantity relates to correlation functions. The argument ‘k’ in both functions specifies the name of the variable that denotes the frequency, which appears as an argument to the Fourier transforms. Thus, our framework has derived the following equations for the model:

$$\begin{aligned}\frac{dq}{dt} &= r - \mu q \\ \frac{dp}{dt} &= 0 \\ \frac{dg}{dt} &= 0,\end{aligned}$$

which are exact because this model is a simple immigration-death model, with no nonlinear processes (and hence no correction to mean-field) and no processes that introduce spatiotemporal correlations.

Adding metapopulation dynamics to this dynamic landscape, we now have two entity types: unoccupied patches (type 1), that can be colonised by occupied patches (type 2) (Fig. 2a, b). The dynamics are now represented by the following generator

$$L = L_1^{\text{IM}}(r) + L_{21}^{\text{I}}(c) + L_{12}^{\text{CT}}(e) + L_1^{\text{D}}(\mu) + L_2^{\text{D}}(\mu),$$

where the last two terms represent patch of either type being destroyed (with the same rate  $\mu$ ), and the first one represents the fact that patches are unoccupied (type 1) when they are created. The second and third terms describe the metapopulation dynamics. The second term represents colonisation, where occupied patches turning an unoccupied patch into an occupied patch with kernel  $c$  (i.e., the colonisation rate is  $c(r)$  if the two patches are separated by distance  $r$ ), and has superscript ‘I’ (denoting ‘Infection’) because this is functionally equivalent to type 2 patches ‘infecting’ type 1 patches. The third term corresponds to spontaneous changes in type (superscript ‘CT’) from type 2 to type 1, representing extinction of occupied patches.

These processes are represented in Model Constructor by the Mathematica process variable

```
processes={Immigration[1,r,1],Infection[2,1,
c,$c,1],
ChangeInType[1,2,e,1],
DensityIndependentDeath[1,mu,1],DensityIndependentDeath[2,mu,1]}
```

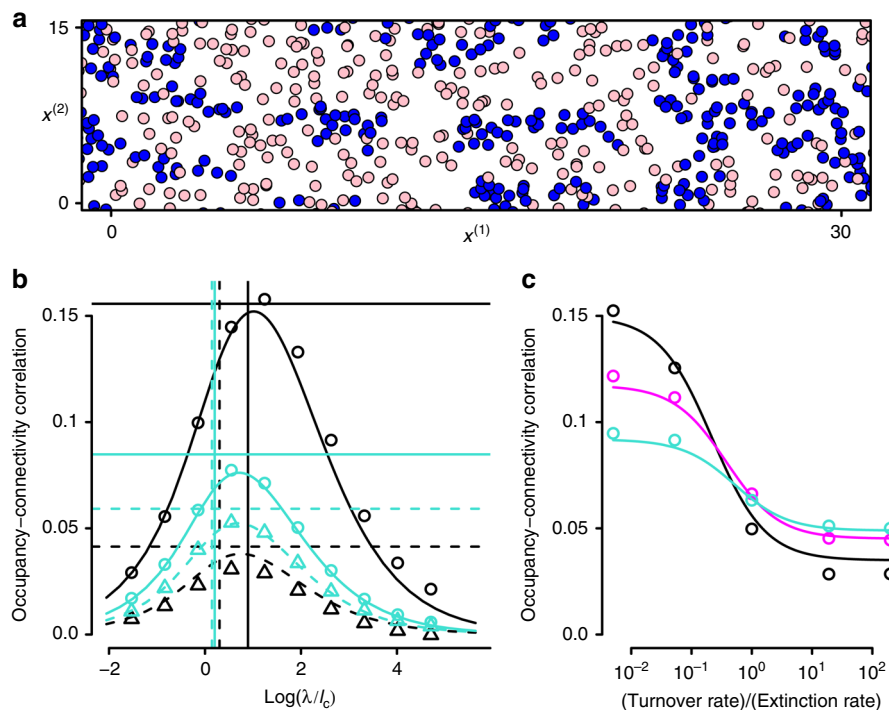
where the argument to the `Infection` process represents the kernel  $c$ , and the argument  $\$c$  represents its Fourier transform. The equations for the quantities  $q,g,p$  can again be obtained using the `HQfALL`, `HGfALL`, `HPfALL` functions. For example,  $dq_1/dt$ ,  $dp_2/dt$  and  $dg_{12}(k)/dt$  are, respectively, given by the outputs of the following commands:

```
HQfALL[{q,p,g},processes,1]
>r - mu q[1] + e q[2] - q[1] q[2] $c[0]
HPfALL[{q,p,g},processes,2,k,2]
>2\pi f_0 k g[1,2,k] $c[k] dk - e p[2] - mu p[2]
+ (p[2] q[1] + p[1] q[2]) $c[0]
HGfALL[{q,p,g},processes,1,2,k]
>-e g[1,2,k] - 2 mu g[1,2,k] + e g[2,2,k]
+ g[1,1,k] q[2] $c[0] - g[1,2,k] q[2] $c[0]
+ g[1,2,k] q[1] $c[k] - g[2,2,k] q[1] $c[k]
- q[1] q[2] $c[k]
```

These equations are identical to those derived in ref. <sup>15</sup> for the case of an uncorrelated landscape (see the Methods section for the full set of differential equations). Our Mathematica toolbox contains tools for readily computing expressions for these quantities at equilibrium, as well as evaluating them numerically (see Supplementary Note 2 section 2.3).

Simulation code is obtained by the ‘Model simulator’ toolbox (Supplementary Note 2 section 2.2). The model is specified using a very similar markup to that for ‘Model Constructor’, except that parameters are given numerical values rather than symbolic names, and kernels have to be assigned specific functional forms. For instance, when colonisation has a top hat kernel with strength  $\tilde{c}(0) = 1.5$  and length scale 3, and the other parameters are  $(r, \mu, e) = (1.1, 0.9, 0.5)$ , the syntax for Model Simulator is specified in file `modelCaseStudy1.txt` containing the text:

```
Immigration[1.1,1]
Infection[2,1,tophat[1.5,3]]
ChangeInType[1,2,0.5]
DensityIndependentDeath[1,0.9]
DensityIndependentDeath[2,0.9]
```



**Fig. 3** Occupancy-connectivity correlation in the dynamic landscape metapopulation. The model is defined in Fig. 2b. Snapshot **a** illustrates that patches are more likely to be occupied (blue symbols) rather than unoccupied (pink symbols) when they are close to other patches **b**. Connectivity-occupancy correlation depends strongly on the ratio of the length scale  $\lambda$  of the connectivity kernel to the length scale  $l_c$  of the colonisation kernel. Lines show analytical predictions and symbols are simulation results. Colonisation and connectivity kernels are chosen to have the same standard deviation when  $\lambda = l_c$ . The highest correlation occurs when  $\lambda$  is 3–4 times larger than  $l_c$ . The optimal length scale depends only weakly on whether the landscape is static (solid lines, circles) or dynamic (dashed lines, triangles), or whether occupancy  $P_0$  is high ( $P_0 = 0.667$ , cyan) or low ( $P_0 = 0.167$ , black). The correlation for the exponential connectivity kernel with the optimal choice of length scale is within 10–20% of the horizontal lines, which represent the occupancy-connectivity correlation for the optimal connectivity kernel shape  $h_*$ . The choice  $\lambda = \lambda_*$  (vertical lines), where the connectivity kernel has the same standard deviation as the optimal connectivity kernel  $h_*$ , gives a higher occupancy-connectivity correlation than  $\lambda = l_c$ . **c** Occupancy-connectivity correlation decreases when rate of patch turnover  $\mu$  increases relative to patch extinction rate  $e$ . Lines are analytical approximation, and symbols are the results of simulations. Parameters  $r$  and  $\mu$  are chosen so that patch density = 1 and mean-field occupancy  $P_0$  is kept constant as  $\mu/e$  changes; colour denotes  $P_0 = 0.167$  (black);  $P_0 = 0.333$  (magenta);  $P_0 = 0.667$  (cyan). Error bars not shown as standard errors are smaller than the plotting symbols. Parameter values and kernels are given in the Methods section

Simulations of the model are run by passing, as a command-line argument, the name of the file containing this model definition to the simulation programme ppsimulator we provide (Supplementary Note 2 section 2.2).

Other kernel shapes are implemented within ppsimulator, but for the sake of computational efficiency kernels that represent interactions between individuals are truncated so as to be zero beyond a threshold distance. For example if ‘tophat[1.5,3]’ were replaced by ‘truncatedGaussian[1.5,3]’ then the colonisation strength would still be  $\tilde{c}(0) = 1.5$ , but the kernel would be a truncated Gaussian (i.e., zero beyond a certain threshold distance, by default three standard deviations and Gaussian otherwise) with standard deviation 3. For more details, see Supplementary Note 2 section 2.2.

**Optimal landscape connectivity.** Our first case study concerns the question in conservation biology of how to identify the most important habitat to conserve. A commonly used metric to assess the value of habitat patch to a network, and aid the design of nature reserves, is ‘connectivity’<sup>16</sup>. A critical open question is how well connectivity predicts patch occupancy—that is, whether connectivity is a good way to identify valuable habitat. Connectivity is expressed as a weighted sum of the proximity to other

habitat patches in the form<sup>16</sup>  $S(x_i) = \sum_j h(x_i - x_j)$ , where  $h$  is a ‘connectivity kernel’, and  $x_i$  and  $x_j$  are positions of the  $i$ th patch and  $j$ th patch, respectively. Commonly,  $h$  is taken to be an exponential function  $h = \exp(-[\text{distance between patches}]/\lambda)$ , where  $\lambda$  is usually chosen to equal the species’ average dispersal distance. While this is inspired from metapopulation ideas, there is no underlying theory for how well connectivity  $S$  predicts habitat occupancy, or whether this kernel is the optimal choice<sup>17</sup>. Our framework allows us to solve this problem and find the best-performing connectivity measure.

We start with the dynamic landscape patch occupancy metapopulation model on a landscape where habitat patches are ephemeral (Figs. 2a, b, c; 3a)<sup>15</sup>. We showed above in ‘Using the framework’ how to use our ‘model constructor’ software to compute the quantities  $q$ ,  $p$  and  $g$  for this model. We used this to derive an analytical expression for the correlation between patch occupancy and connectivity  $S$  for a general connectivity kernel  $h$  (see the Methods section), which can be expressed in terms of spatial moments and therefore in terms of our quantities  $q$ ,  $p$  and  $g$ . Considering first an exponential connectivity kernel, we find that the standard choice of the length scale  $\lambda$  is not optimal as it gives a much lower correlation than when  $\lambda$  is 3–4 times larger (Fig. 3b). We further derived an expression for the optimal connectivity kernel, i.e., the one that maximises the

connectivity–occupancy correlation:

$$\tilde{h}_*(\omega) = \frac{\tilde{c}(\omega)}{\tilde{c}(0)(\tilde{c}(0)r + \mu^2) - \mu(e + \mu)\tilde{c}(\omega)}.$$

Here and throughout this paper, a tilde represents Fourier transform;  $\tilde{h}_*(\omega)$  is the Fourier transform of the optimal connectivity kernel,  $\tilde{c}(\omega)$  is the Fourier transform of the colonisation kernel and  $\omega$  is a frequency (the Fourier conjugate variable to space). The parameters  $r$ ,  $\mu$  and  $e$  are defined in the caption of Fig. 2 and have the same meanings as in ‘Using the framework’ above (i.e., rates of patch creation, destruction and extinction respectively). It is clear from this expression that, in general, this kernel will have a different shape from either the exponential function or the colonisation kernel. However, we find that an exponential connectivity kernel with the optimal choice of  $\lambda$  performs nearly as well (Fig. 3b). Thus, we show that an exponential connectivity kernel is a reasonable choice for landscape design, but we suggest that its ability to identify important habitat would be improved if its standard deviation were set to match that of the optimal kernel (Fig. 3).

This means that the connectivity kernel should not just depend on the dispersal properties as is usually assumed<sup>16</sup>, but also on the metapopulation occupancy and landscape turnover as well (Fig. 3b). Moreover, we find that connectivity–occupancy correlation becomes weaker when the landscape becomes more dynamic (i.e., patches are created and destroyed more rapidly, Fig. 3c)<sup>18</sup>. Also, for static landscapes (left side of Fig. 3c), the correlation is stronger when the mean patch occupancy is lower, but this trend is reversed for dynamic landscapes (right side of Fig. 3c). Thus, connectivity is of less use to biological conservation when either occupancy or habitat turnover is high<sup>19</sup>.

**Genetic similarity.** Our second case study is in the field of population genetics. Molecular ecologists study patterns of genetic similarity and differentiation between populations or across space to infer the underlying processes that cause and maintain that genetic structure. A long-standing problem is to find tractable models where local density is regulated so that population dynamics are stable<sup>20</sup>. Many ingenious solutions to this problem include imposing constant local density artificially<sup>21–23</sup>. However, since the aim of these models is to relate pattern to process, a much preferable solution would be to do this via local density-dependent population dynamics. Such models are regarded as too ‘hopelessly complicated’<sup>24,25</sup> to study using standard methods: even in the simplest models, it is highly nontrivial to obtain analytical expressions for genetic similarity, and the incorporation of more realistic processes (e.g., such as selection) poses even greater difficulty<sup>24,25</sup>. Our framework overcomes these obstacles, allowing us to consider local density-dependent population regulation explicitly, and to derive analytical expressions for how genetic similarity varies in space. We illustrate this with an example including local competition and limited dispersal, selection and mutation between alleles in neutral and non-neutral loci, and also adaptation to heterogeneous environments, which is important for example for tropical forests<sup>22</sup>.

The model involves two habitat types, and haploid individuals with four different genotypes (one selective and one neutral locus, two alleles at each). Habitat patches appear independently as a Poisson process with rate  $\kappa$  and vanish with rate  $\mu$ . All individuals have density-independent mortality (rate  $m$ ) and density-dependent mortality with the same interaction kernel  $c$  between all genotypes. Birth is density-independent and offspring are distributed relative to their parents with kernel  $d$ , but the fecundity depends on habitat availability (via the patch kernel  $r$ )

and the match between the habitat type and the genotype of the individual via the function  $\phi^r = f_0 \cdot (1 \pm \tau)$ , where  $f_0$  is a constant base fecundity and  $\tau$  is the strength of selection. Therefore, each habitat patch a distance  $x$  away from the individual contributes  $r(x) \cdot f_0 \cdot (1 + \tau)$  to the fecundity of that individual if allele in selective locus matches the type of habitat, and  $r(x) \cdot f_0 \cdot (1 - \tau)$  otherwise. Offspring inherits their parent’s genotype (i.e., reproduction is clonal) unless there is a mutation, which happens independently at each allele at rate  $\nu$  per birth and specified by the function  $\psi^\nu$ . All model components are shown in Fig. 4a, and a typical snapshot of the dynamics is illustrated in Fig. 4b.

We are interested in the similarity functions  $F_n(x)$  and  $F_s(x)$  at the neutral and selective loci, respectively, i.e., the probability that two individuals separated by distance  $x$  have the same allele at the locus in question. These can be computed from the probability densities to find specific individuals at this spatial separation. Using the following function

$$F_i(x) = \frac{\text{probability density for individuals at } y \text{ and } x+y \text{ with allele } i}{\text{probability density for any two individuals at } y \text{ and } x+y},$$

where  $i = 1, 2$  denote alleles in neutral locus, and  $i = A, B$  denote alleles in selective locus, the similarity functions become  $F_n(x) = F_1(x) + F_2(x)$ ,  $F_s(x) = F_A(x) + F_B(x)$ . The probability densities in the equation above can be expressed in terms of the two-point correlation functions (second spatial moments)  $k_{ij}^{(2)}(x_i, x_j)$  that in turn can be expressed in terms of the quantities  $q$ ,  $p$  and  $g$  in the perturbation expansion<sup>11</sup>, i.e.,  $k_{ij}^{(2)}(x) = q_i q_j + \epsilon^d (g_{ij}(x) + q_i p_j + p_i q_j) + o(\epsilon^d)$ .

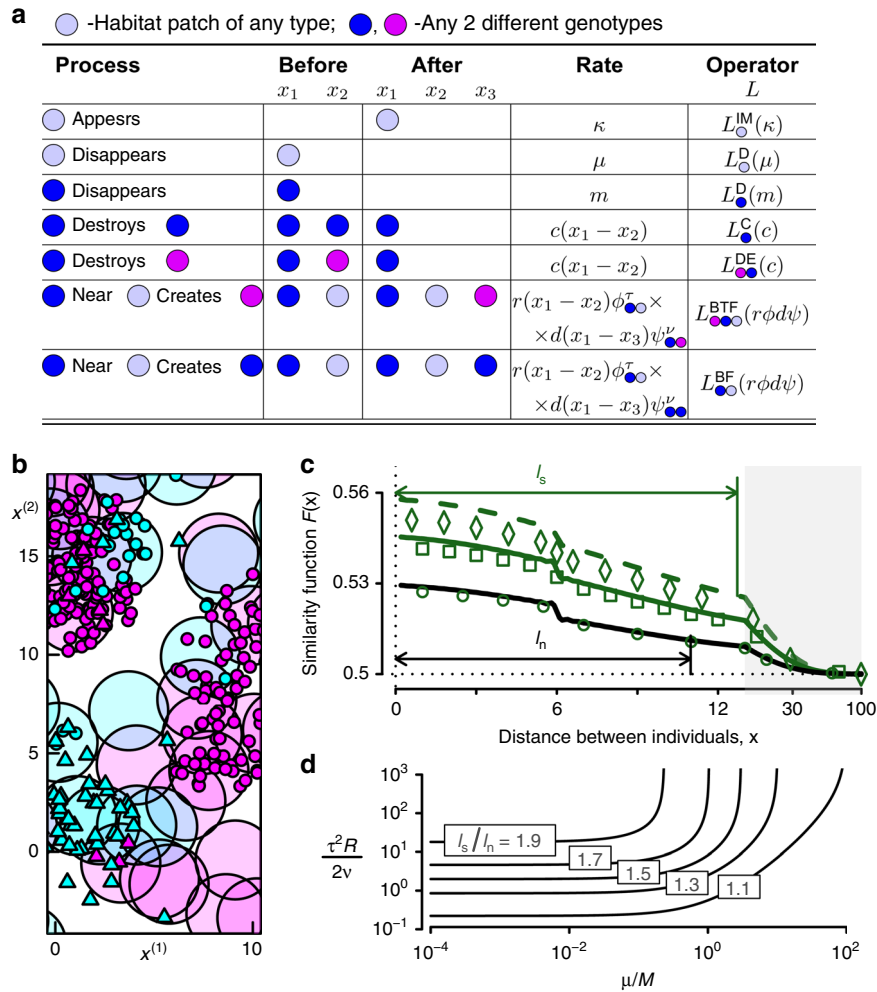
Using our framework, we obtained analytical expressions for how genetic similarity depends on the distance between individuals (Methods) which show good agreement with numerical simulations, Fig. 4c. Genetic similarity at the selective locus is always higher than at the neutral locus, and also extends further in space due to having different length scales  $l_n$  and  $l_s$  (Fig. 4c), where  $l_s$ ,  $l_n$  are, respectively, the length scales of genetic similarity at the selective and neutral loci. Our analytical expressions for genetic similarity (shown in the Methods section) enable us to understand what controls the differences in these length scales. The results become particularly simple under the plausible assumption that mutation is rare ( $\nu \ll 1$ , see the Methods section):

$$\frac{l_s}{l_n} = \frac{1 + \frac{r^2 R}{2\nu} \frac{M}{\mu} (1 - r^2)}{1 + \frac{r^2 R}{2\nu} \frac{M}{\mu} (1 - r)},$$

where  $R = \frac{N_e}{2\phi_0}$  is the ratio of the total density of individuals  $N_e$  to the total density of habitat patches,  $M = 4\nu q_h f_0 \tilde{r}(0) \tilde{d}(0)$  is the per-capita creation rate of mutants in the population,  $\mu$  is the turnover rate of habitat patches and  $r = (1 + \mu/M)^{-1/2}$ . We see that the ratio of these length scales is controlled by just two parameter combinations:  $\mu/M$ , which quantifies how ephemeral the landscape is relative to the appearance of mutants, and  $\frac{r^2 R}{2\nu}$ , which will be numerically large when  $\nu \ll 1$  unless selection is very weak or habitat patches vastly outnumber the organism. Similarity at the selective locus will resemble that at the neutral locus if mutation is common or selection weak, but will extend up to twice as far when habitat turnover is slow relative to mutation (Fig. 4d).

**Optimal foraging.** Our third case study originates from the field of movement ecology, in which a long-standing challenge is to understand the causes and consequences of movement behaviour on foraging efficiency<sup>26,27</sup>. In environments with patchily





**Fig. 4** Spatial patterns of genetic differentiation in heterogeneous environments. **a** Graphical definition of the population model. **b** A snapshot of typical dynamics of the model: two types of habitat (light cyan or magenta discs), four genotypes of individuals (small symbols, colour denotes the preferred habitat of allele at selective locus, shape—triangle or circle—denotes allele at neutral locus), see the Methods section for values of parameters. **c** Similarity function (probability of allelic identity by state) at neutral (black) and selective (green) loci found theoretically (lines) and by simulations (symbols) have different length scales  $l_n$  and  $l_s$  (defined from the first moment of the similarity function). Parameters ( $\tau$  = selection strength,  $\mu$  = habitat turnover) equal (1, 0.5) for diamonds and the dashed line, (1, 1) for squares and solid green line, and (0.1, 1) for circles and solid black line. Here, at  $x > 13$  a log-scale is used (grey area). **d** When mutation is rare ( $\nu \ll 1$ ), the factor  $l_s/l_n$  by which genetic similarity extends further at the selective locus than at the neutral locus depends only on the two parameter combinations  $\frac{\tau^2 R}{2\nu}$  (with  $R = \frac{N_e}{2q_h}$ ) and  $\frac{\mu}{M}$ . The length scales are most different when selection is strong, mutation rare, habitat static, and there are many individuals per habitat patch. Error bars not shown as standard errors are smaller than the plotting symbols. Parameter values are given in the Methods section

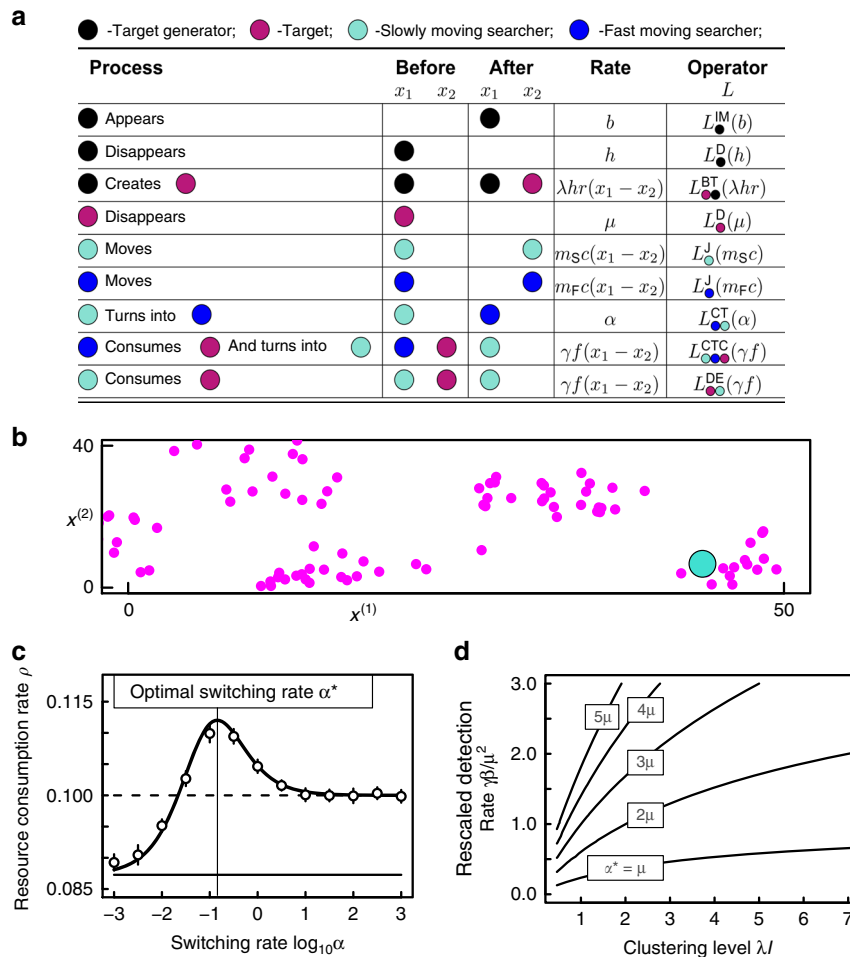
distributed resources, commonly observed behaviours involve slow foraging movements within patches, interrupted by fast exploratory movements between the patches<sup>28</sup>. Individual-based models are needed to account for stochasticity and resource heterogeneity in foraging<sup>29</sup>, but there are several processes at work so it is very difficult from simulations to understand the key evolutionary drivers and ecological consequences of variation in foraging behaviour. Our framework solves this problem by providing simple analytical expressions that quantify the ecological factors determining the optimal foraging strategy.

We used the graphical model description (Fig. 5a) to construct a model in which aggregates of resources are continuously generated at new locations, while existing resource units decay, resulting from the consumer's point of view in an unpredictable resource distribution. The model consists of resources (targets), which are generated in clusters as follows. Target 'generators' appear as a Poisson process with intensity  $b$ , and disappear at rate  $h$ . During their lifetime, target generators create targets at rate  $h\lambda$

at a distance from the generator determined by kernel  $r$  (normalised so  $\tilde{r}(0) = 1$ ). We assume  $h \rightarrow \infty$  such that  $h\lambda$  is constant, so that each generator instantaneously creates a Poisson distributed number of targets with mean  $\sim \lambda$ . Targets vanish at rate  $\mu$ . The forager moves as a jump process with jump kernel  $c$  at rates  $m_S$  and  $m_F$ , respectively, when in slow or fast mode. Foragers consume targets at rate  $\gamma$  with kernel  $f$ . We assume the consumer has evolved to switch to a fast movement mode (at a rate we denote by  $\alpha$ ) when it does not encounter new resources, and to switch back to slow movement mode after encountering a resource unit (Fig. 5a, b).

Our mathematical formalism yields a simple closed-form expression for the mean consumption rate (Fig. 5c),  $\rho \approx \rho_0 + \epsilon^d \rho_1$ , where  $\rho_0 = \gamma\beta/\mu$  is the mean-field consumption rate,  $\beta = \lambda b$ , and the first-order correction to  $\rho$  is

$$\rho_1 = \frac{\beta\gamma^2(\alpha\lambda\mu I_1 - \beta\gamma I_2)}{\mu(\beta\gamma + \alpha\mu)(\alpha + \mu)}, \quad (1)$$



**Fig. 5** Optimal foraging model. **a** shows a graphical definition of the model and **b** shows a simulation snapshot, illustrating a forager (large cyan circle) that has found a resource aggregate, and has switched to the slow movement mode to consume resources (small magenta circles). **c** shows how mean resource consumption rate  $\rho$  varies from a totally stationary searcher ( $\alpha = 0$ ; continuous horizontal line) through a searcher following the optimal rate  $\alpha^*$  (vertical line) to a highly mobile searcher ( $\alpha \rightarrow \infty$ , dashed line). In **c**, the dots with error bars show simulation results, whereas the lines follow the analytical prediction of  $\rho \approx \rho_0 + \epsilon^d \rho_1$  (see Eq. (1)). The contour lines in **d** show the optimal switching rate  $\alpha^*$  as a function of the rescaled detection rate  $\gamma\beta/\mu^2$  of the searcher and the clustering level  $\lambda I$  of the targets. Error bars represent three standard errors. Parameters are shown in the Methods section

with  $I_1 = \int \tilde{f}^2(\omega) \tilde{r}^2(\omega) d\omega$  and  $I_2 = \int \tilde{f}^2(\omega) d\omega$ . This expression behaves non-monotonically as a function of the switching rate, and is maximised at an intermediate value  $\alpha^*$  (vertical line in Fig. 5c). This is because a consumer that remains continuously in the slow mode is not efficient in finding new resource aggregates, whereas a consumer that remains continuously in the fast mode misses the opportunity of elevated resource availability within aggregates.

We also find a simple expression for the optimal switching rate:  $\alpha^* = \mu f\left(\frac{\rho_0}{\mu}, \frac{1}{\lambda I}\right)$ , where  $f(x, y) = xy + \sqrt{x(1+xy)(1+y)}$ ,  $\lambda$  is the mean number of resource items per cluster, and  $I = I_1/I_2$  ( $0 < I < 1$ ) quantifies the spatial scale of the resource detection process relative to the resource cluster size. Thus, although we started with a model with nine processes, the optimal foraging strategy is only determined by the two parameter combinations  $\frac{\rho_0}{\mu}$  and  $\lambda I$  (Fig. 5d)—for example, an increase in resource acquisition rate has the same effect on the switching rate as the same proportionate increase in resource production rate. It therefore pays to spend a long time in the slow foraging mode (small  $\alpha^*$ ) if the resources are highly aggregated (large  $\lambda$ ) or if the consumer's resource detection efficiency ( $\gamma$ ) is low, Fig. 5d. However, if fast movement requires more energy than the slow mode, we find that

$\alpha^* = 0$  when these energetic costs per unit time exceed a threshold (see the Methods section). This threshold value shows that it may be better to stay put and wait for resources to be generated locally, rather than to roam widely, when the resource consumption rate ( $\rho_0$ ) is low, resource clusters contain few items ( $\lambda$  small), or the spatial scale of clusters and/or resource detection is large.

## Discussion

Our framework is game-changing because (i) it makes mathematical approximations available for a very wide class of individual-based models, and (ii) this approximation is available to the non-specialist because it requires no laborious derivations on the part of the researcher. Since we have derived expressions for general reactant-catalyst-product model, and implemented these in Mathematica, a researcher interested in a particular model has only to convert their model into the markup understood by our software (see the Methods section and Fig. 2a–c) and the approximation is generated for them automatically. Without our method, mathematical results for a new model would only be available to a researcher with the skill and patience to spend weeks performing algebra without error. Our software also generates code for simulating the same model. The mathematical

validity of our approach is proved in Supplementary Note 1, and the validity of the approximation in these examples is shown by its close agreement with simulations in Figs. 3–5. Since our framework is based on a perturbation expansion, it is guaranteed to give a good approximation when the length scale of interaction kernels is large enough compared to the separation between individuals in the system<sup>15</sup>. One way to assess the accuracy of the approximation is to compare  $p$  to  $q$ : if the first-order correction  $p$  is comparable in magnitude to the mean-field density  $q$ , then approximation may not be very accurate for those parameter values because higher order corrections are likely to be important.

Our examples illustrate the conceptual and methodological advantages of analytical methods over simulations. In the first case study, we have a general expression for the best-performing connectivity kernel, valid for any colonisation kernel, whereas simulations can only explore specified examples and functional forms. In the third case study, we obtained directly an expression for the optimal foraging strategy, whereas in a simulation study the optimal strategy would have to be searched for separately for each parameter combination. We found in the second and third examples that the behaviour depends on a small number of parameter combinations, which effectively reduces the dimension of the parameter space of the problem. This means that the model behaviour can be explored more economically, and explained more intuitively, than in a simulation study in which each set of parameter values needs to be investigated independently. While simulations will always play an important role in studying spatial and stochastic models, our mathematical framework will facilitate much deeper understanding by making analytical results readily available for models that were previously thought to be too complex.

There are already platforms that allow simulation of spatial, stochastic individual-based models specified in a user-friendly way<sup>10,30</sup>, but ours is the first framework that automatically performs mathematical analyses as well. We have formulated, analysed mathematically and written computer code for the general model containing any combination of reactant–catalyst–product processes. This general model can be used to describe a very large array of processes, so that a wide variety of models of interest in ecology and evolution are straightforward to construct and analyse using the unified framework presented here. Thus, our framework provides a long-sought analytical tool<sup>8</sup> with which many classical questions in ecology, evolution and other fields can be revisited to understand the role of space and stochasticity.

## Methods

**Derivation of the analytical framework.** Here, we give a summary of the approach and key results of our analytical framework; full details of the calculations are given in the Supplementary Notes.

We begin by classifying the participants in demographic processes into three types of individuals: (i) reactants (that are destroyed by the process); (ii) products (that are created by the process) and (iii) catalysts (that are unaffected by the process but whose presence affects the rate at which it occurs). A given demographic process is therefore characterised by integer numbers denoted by  $P$ ,  $R$  and  $C$  that give number of agents in the groups ‘reactants’, ‘products’ and ‘catalysts’. The group ‘products’ is characterised by the set  $\mathbf{P}$  consisting of pairs of the type of agent  $i$  and the location  $x$  (from  $d$ -dimensional space  $x \in \mathbb{R}^d$ ) assigned to the agent:  $\mathbf{P} = \{\{i_1, x_{i_1}\}, \{i_2, x_{i_2}\}, \dots, \{i_p, x_{i_p}\}\}$ . Similarly, reactants are described by  $\mathbf{R} = \{\{j_1, x_{j_1}\}, \{j_2, x_{j_2}\}, \dots, \{j_r, x_{j_r}\}\}$ , and catalysts are described by  $\mathbf{C} = \{\{k_1, x_{k_1}\}, \{k_2, x_{k_2}\}, \dots, \{k_c, x_{k_c}\}\}$ . We introduce the function  $r(\mathbf{P}, \mathbf{R}, \mathbf{C})$ , which depends on types of agents and on distances between locations of agents used in  $\mathbf{P}$ ,  $\mathbf{R}$  and  $\mathbf{C}$ . The function  $r$  determines interactions between agents by specifying those coordinates and indices which are the same in  $\mathbf{P}$ ,  $\mathbf{R}$ ,  $\mathbf{C}$ . For example, the ‘Change in type’ process where agents of type  $j_1$  transform with rate  $r_0$  into agents of type  $i_1$  without changing their locations is described by:  $\mathbf{P} = \{\{i_1, x_{i_1}\}\}$ ,  $\mathbf{R} = \{\{j_1, x_{j_1}\}\}$ ,  $\mathbf{C} = \emptyset$ , and  $r(\mathbf{P}, \mathbf{R}, \mathbf{C}) = r_0 \delta(x_{i_1} - x_{j_1})$ .

For the purpose of this paper, it is convenient to work with cumulants  $u$  instead of moments, as cumulants carry the same information as moments but lead to a more simple perturbation expansion<sup>11</sup>. We define the operator  $Q^\Delta$  that determines

the full hierarchy of equations for cumulants:

$$\frac{\partial}{\partial t} u(\eta, t) = (Q^\Delta u)(\eta, t). \quad (2)$$

Here, all orders of cumulants are denoted by  $u(\eta, t)$ , where  $\eta$  denotes any finite number of points in  $d$ -dimensional continuous space  $\mathbb{R}^d$  occupied by agents of a specified type. For example, for a one-point configuration for an agent of type  $m$  one has  $\eta = \{x, m\}$ , and the expression  $u(\eta, t)$  gives a first-order cumulant  $u_m^{(1)}(x, t)$  which is the density of agents  $m$  at location  $x$ , denoted as  $D_m(x, t) = u_m^{(1)}(x, t)$ . For a two-point configuration  $\eta = \{x, m; y, n\}$  one obtains a second cumulant  $u(\eta, t) = u_{m,n}^{(2)}(x, y, t)$  that determines spatial covariance denoted  $\text{Cov}_{m,n}(x, y, t) = u_{m,n}^{(2)}(x, y, t)$ .

Following ref. <sup>11</sup>, we made the following steps in the derivation (for detailed derivations see Supplementary Note 1): (1) using notions of locally finite configurations, transform an agent-based model description into the dynamics of spatial moments; (2) using the correspondence between correlation functions and cumulants, obtain operator  $Q^\Delta$  for the evolution of cumulants. We obtained:

$$\begin{aligned} (Q^\Delta u)(\eta, t) = & \left( \int_{\mathbb{R}^{d(P+R+C)}} dx_{i_1} \dots dx_{i_p} dx_{j_1} \dots dx_{j_r} dx_{k_1} \dots dx_{k_c} r(\mathbf{P}, \mathbf{R}, \mathbf{C}) \right. \\ & \times \prod_{k=k_1}^{k_c} \left( 1 + D_{x_k}^{(i(k))} \right) \left[ \prod_{i=i_1}^{i_p} (1 + D_{x_i}^{(i(i))}) - \prod_{j=j_1}^{j_r} (1 + D_{x_j}^{(j(j))}) \right] V \Big) (\eta, t), \end{aligned} \quad (3)$$

where we used the following operations  $D_{x_m}^{(m)}$  and  $D_{x_m}^{(i(m))}$ , that can be considered as related to creation and annihilation operators:

$$\begin{aligned} (D_{x_m}^{(m)} u)(\eta) &:= u(\eta \cup x_m^{(m)}), \\ (D_{x_m}^{(i(m))} u)(\eta) &:= \sum_{y_m \in \eta} \delta(x_m^{(m)} - y_m^{(m)}) u(\eta \setminus x_m^{(m)}), \end{aligned}$$

here  $x_m^{(m)}$  denotes an agent of type  $m$  at location  $x_m$ . In Eq. (3),  $V$  can be thought of as a generating function:

$$V := (\exp^{*-1} u) * \left( D_{x_{j_1}}^{(j_1)} \dots D_{x_{j_r}}^{(j_r)} D_{x_{k_1}}^{(k_1)} \dots D_{x_{k_c}}^{(k_c)} \exp^* u \right),$$

where we used the operation  $*$  that is defined below, also see ref. <sup>11</sup>. First, let  $\Gamma_0$  denotes the set of all finite subsets  $\eta$  of  $\mathbb{R}^d$ . For any functions  $u, v$  on  $\Gamma_0$ , the  $*$  operation is defined as:

$$(u * v)(\eta) := \sum_{\eta_1 \sqcup \eta_2 = \eta} u(\eta_1) v(\eta_2),$$

where the symbol  $\sqcup$  denotes a disjoint union defined as

$$\sum_{\eta_1 \sqcup \eta_2 = \eta} u(\eta_1) v(\eta_2) = \sum_{\xi \subset \eta} u(\xi) v(\eta \setminus \xi),$$

thus, the symbol  $*$  defines a convolution. Using these notations, the function  $(\exp^* u)(\eta)$  is

$$(\exp^* u)(\eta) := \sum_{n=0}^{\infty} \frac{1}{n!} u^{*n}(\eta) = 1^*(\eta) + \sum_{n=1}^{\infty} \frac{1}{n!} \sum_{\eta_1 \sqcup \dots \sqcup \eta_n = \eta} u(\eta_1) \dots u(\eta_n),$$

where

$$1^*(\eta) = 0^{|\eta|} = 1_{|\eta|=0} = \begin{cases} 1, & \eta = \emptyset; \\ 0, & \eta \neq \emptyset. \end{cases}$$

The function  $\exp^{*-1} u$  is the inverse with respect to the  $*$  convolution, i.e., as was shown in e.g., <sup>31</sup>, for any  $u$  with  $u(\emptyset) = 0$ , there exist a function  $\exp^{*-1} u$  such that  $(\exp^{*-1} u) * (\exp^* u) = 1^*$ ,

$$(\exp^{*-1} u)(\eta) := \sum_{n=0}^{\infty} (-1)^n ((\exp^* u) - 1^*)^{*n}(\eta).$$

The system of Eq. (2) cannot be solved exactly, because an equation for a cumulant of any order depends on higher order cumulants. In our derivations we follow the method from<sup>11</sup>, which assumes that interaction between individuals is long ranged; details of derivations are presented in Supplementary Note 1, sections 1.1–1.3. Using this exact expression of the full hierarchy of moment equations for the general model, we apply the perturbation scheme<sup>11,13</sup> to derive a controlled approximation that gives asymptotically exact results when agents interact over large enough scales. It is assumed that spatial interactions between individuals depend on their separation  $x$  (in  $d$ -dimensional space) according to interaction ‘kernels’ of the form  $\epsilon^d f(\epsilon x)$  where  $1/\epsilon$  is the typical length scale of interactions. In models with long-ranged interactions (i.e., where typical scale of interaction is large), the parameter  $\epsilon$  is a small parameter that can be used to develop a perturbation expansion. It is this small parameter  $\epsilon$  that makes this approximation controlled: the smaller  $\epsilon$ , the better leading terms in a perturbation expansion describe the exact solution.

We showed mathematically that for a model made up of a set general process defined by  $\mathbf{P}$ ,  $\mathbf{R}$ ,  $\mathbf{C}$  and the interaction function  $r_\kappa(\mathbf{P}, \mathbf{R}, \mathbf{C})$ , the mean densities and spatial covariance (including autocovariance) satisfy the following expansion:

$$\text{density of species } m = q_m(\epsilon x, t) + \epsilon^d p_m(\epsilon x, t) + o(\epsilon^d),$$

$$\text{spatial covariance between species } m \text{ and } n = \epsilon^d g_{m,n}(\epsilon x_1, \epsilon x_2, t) + o(\epsilon^d),$$

where

$$\begin{aligned} \frac{dq_m(x, t)}{dt} &= \sum_{\kappa \in \text{model components}} H_{q_m}^{(\kappa)}(x, t) \\ \frac{dp_m(x, t)}{dt} &= \sum_{\kappa \in \text{model components}} H_{p_m}^{(\kappa)}(x, t) \\ \frac{dg_{m,n}(x_1, x_2, t)}{dt} &= \sum_{\kappa \in \text{model components}} H_{g_{m,n}}^{(\kappa)}(x_1, x_2, t) \end{aligned}$$

For a general (arbitrary) product–reactant–catalyst process  $\kappa$ , we derived the corresponding contributions  $H_{q_m}^{(\kappa)}$ ,  $H_{p_m}^{(\kappa)}$  and  $H_{g_{m,n}}^{(\kappa)}$  to the differential equations for  $q_m$ ,  $p_m$  and  $g_{m,n}$ , see Eq. (2) in Fig. 1e. The expression  $H_{q_m}^{(\kappa)}$  for the mean-field density  $q_m(x, t)$  of an agent of type  $m$  for an arbitrary spatially heterogeneous product–reactant–catalyst process is shown below:

$$\begin{aligned} H_{q_m}^{(\kappa)}(x, t) &= \int_{\mathbb{R}^{d(P+R+C)}} dx_1 \dots dx_p dx_{j_1} \dots dx_r dx_{k_1} \dots dx_c r_\kappa(\mathbf{P}, \mathbf{R}, \mathbf{C}) \\ &\times \left( \sum_{i=1}^p \delta_m \delta(x_i - x) - \sum_{j=1}^r \delta_m \delta(x_j - x) \right) \prod_{\beta_i \in \{j_1, \dots, k_c\}} q_{\beta_i}(x_{\beta_i}, t). \end{aligned} \quad (4)$$

Expression for  $H_{p_m}^{(\kappa)}(x, t)$  and  $H_{g_{m,n}}^{(\kappa)}(x_1, x_2, t)$  are shown below, while their derivations are presented in Supplementary Note 1.

$$\begin{aligned} H_{p_m}^{(\kappa)}(x, t) &= \int_{\mathbb{R}^{d(P+R+C)}} dx_1 \dots dx_p dx_{j_1} \dots dx_r dx_{k_1} \dots dx_c r_\kappa(\mathbf{P}, \mathbf{R}, \mathbf{C}) \\ &\times \left( \sum_{i=1}^p \delta_m \delta(x_i - x) - \sum_{j=1}^r \delta_m \delta(x_j - x) \right) \\ &\times \left( \sum_{\beta_i \in \{j_1, \dots, k_c\}} p_{\beta_i}(x_{\beta_i}, t) \prod_{\beta_i \in \{j_1, \dots, k_c\} \setminus \beta_i} q_{\beta_i}(x_{\beta_i}, t) \right. \\ &\left. + \sum_{\beta_1 \neq j_1}^{k_c} \sum_{\beta_2 \neq j_1+1}^{k_c} g_{\beta_1 \beta_2}(x_{\beta_1}, x_{\beta_2}, t) \prod_{\beta_i \in \{j_1, \dots, k_c\} \setminus \beta_1 \setminus \beta_2} q_{\beta_i}(x_{\beta_i}, t) \right). \end{aligned} \quad (5)$$

$$\begin{aligned} H_{g_{m,n}}^{(\kappa)}(x_1, x_2, t) &= \int_{\mathbb{R}^{d(P+R+C)}} dx_1 \dots dx_p dx_{j_1} \dots dx_r dx_{k_1} \dots dx_c r_\kappa(\mathbf{P}, \mathbf{R}, \mathbf{C}) \\ &\times \left\{ \sum_{i=1}^p \sum_{i'=1}^p (\delta_m \delta_{m'} \delta(x_i - x_1) \delta(x_{i'} - x_2) + \delta_{m'} \delta_m \delta(x_i - x_2) \delta(x_{i'} - x_1)) \right. \\ &- \sum_{j=1}^r \sum_{j'=1}^r (\delta_m \delta_{m'} \delta(x_j - x_1) \delta(x_{j'} - x_2) + \delta_{m'} \delta_m \delta(x_j - x_2) \delta(x_{j'} - x_1)) \\ &+ \sum_{k=1}^{k_c} \delta_{nk} \delta(x_k - x_1) \sum_{i=1}^p \delta_m \delta(x_i - x_2) + \sum_{i=1}^p \delta_m \delta(x_i - x_1) \sum_{k=1}^{k_c} \delta_{nk} \delta(x_k - x_2) \\ &- \sum_{k=1}^{k_c} \delta_{mk} \delta(x_k - x_1) \sum_{j=1}^r \delta_n \delta(x_j - x_2) - \sum_{j=1}^r \delta_n \delta(x_j - x_1) \sum_{k=1}^{k_c} \delta_{mk} \delta(x_k - x_2) \\ &\times q_{j_1}(x_{j_1}, t) \dots q_{j_r}(x_{j_r}, t) q_{k_1}(x_{k_1}, t) \dots q_{k_c}(x_{k_c}, t) \\ &+ \sum_{i=1}^p \sum_{\beta_i \in \{j_1, \dots, k_c\}} [\delta_m \delta(x_i - x_1) g_{\beta_i n}(x_{\beta_i}, x_2) + \delta_n \delta(x_i - x_2) g_{m \beta_i}(x_1, x_{\beta_i})] \\ &\times \prod_{\beta_i \in \{j_1, \dots, k_c\} \setminus \beta_i} q_{\beta_i}(x_{\beta_i}, t) \\ &- \sum_{j=1}^r \sum_{\beta_j \in \{j_1, \dots, k_c\}} [\delta_m \delta(x_j - x_1) g_{\beta_j n}(x_{\beta_j}, x_2) + \delta_n \delta(x_j - x_2) g_{m \beta_j}(x_1, x_{\beta_j})] \\ &\times \prod_{\beta_j \in \{j_1, \dots, k_c\} \setminus \beta_j} q_{\beta_j}(x_{\beta_j}, t) \Big\}. \end{aligned} \quad (6)$$

Using expressions  $H_{q_m}^{(\kappa)}(x, t)$ ,  $H_{p_m}^{(\kappa)}(x, t)$  and  $H_{g_{m,n}}^{(\kappa)}(x_1, x_2, t)$ , it is straightforward to show that values of quantities  $q_m(\epsilon x, t)$ ,  $\epsilon^d p_m(\epsilon x, t)$  and  $\epsilon^d g_{m,n}(\epsilon x_1, \epsilon x_2, t)$  calculated using the given kernels  $a$  are identical to the values of quantities  $q_m(x, t)$ ,  $p_m(x, t)$  and  $g_{m,n}(x_1, x_2, t)$  calculated using the scaled kernels  $a_\epsilon$ ,  $a_\epsilon(x - y) = \epsilon^d a(\epsilon(x - y))$ .

**Computer code.** The analytical software ‘Model Constructor’ is developed in Mathematica<sup>32</sup>. The code first requires a user to specify the model in terms of products, reactants and catalysts, and the interactions. Then, using Fourier transform and a simplifying assumption of translational invariance, the code uses the input data to obtain the analytical expressions for  $H_{q_m}^{(\kappa)}(x, t)$ ,  $H_{p_m}^{(\kappa)}(x, t)$  and  $H_{g_{m,n}}^{(\kappa)}(x_1, x_2, t)$  for a given system in 1D, 2D or 3D infinite space. Also, the code can write down these analytical expressions in real space and in Fourier space into .tex file. The analytical software has been checked to reproduce the results for 15 specific processes (see Supplementary Notes) that were derived analytically.

The expressions produced by Model Constructor assume translational invariance, i.e., that the agents do not interact with a spatially heterogeneous

extrinsic environment, and the expressions are averaged over initial conditions that are translationally invariant. This is not due to a limitation of the underlying mathematical framework, as Eqs. (4), (5), (6) do not make this assumption. However, in the absence of translational invariance the equations for  $g$  and  $p$  involve two spatial co-ordinates rather than one, and as a result are much more challenging to solve both analytically and numerically. Any analytical solution would require exploiting whatever symmetries are present in the initial condition, which depends on the details of each individual case. While we are studying some specific cases without translational invariance (work in progress), software that allows spatial heterogeneity to be specified in a general way is beyond the scope of this paper.

The simulation software ‘Model simulator’ is a C-programme for simulating continuous-time point processes. Each point is associated with a coordinate and a discrete species attribute. Points are located either on 1D or 2D torus space. The user defines the set of processes and the initial configuration after which the simulator runs the Gillespie algorithm<sup>33</sup> in such a way that the information of point locations are taken into account, i.e., the system is not assumed to be well-mixed. The state of the configuration can be outputted at user-defined constant time intervals. Auxiliary R-functions are provided for calculating summary statistics and creating figures and animations based on the simulation. Input for the simulator is given by means of text files and few command line arguments. Output of the simulator is written in text files. The simulation software has been checked to replicate the results that we obtained earlier using more specific implementations (e.g. ref. 11) that were fully independent of the current implementation (e.g. coded in different programming languages).

In all systems studied in this paper results of both types of software, analytical and simulation, match each other in the way that the mathematical theory suggests. The detailed derivations of underlying mathematical expressions, and the detailed tutorials for analytical and simulation software are shown in Supplementary Notes.

**Case studies.** Model Constructor was used to obtain the analytical results used in the case studies, and for representative parameter values these were compared with simulation results obtained using Model Simulator. The toolboxes and files containing the markup for the case studies are included at Figshare, <https://doi.org/10.6084/m9.figshare.9633161>. We give here the essential details for obtaining the case study results used in the paper; further details are given in Supplementary Notes 3, 4 and 5.

**Optimal landscape connectivity model.** For case study 1, the model is defined verbally in Fig. 2a and the processes are listed in Fig. 2b.

Perturbation equations: The section ‘Using the framework’ in the main paper contains a full description of the steps required to derive the following expressions for  $q_i$  and  $g_{ij}$  ( $p_i$  is not needed for the rest of the calculation):

$$\begin{aligned} \frac{dq_1}{dt} &= r - \mu q_1 + e q_2 - q_1 q_2 \tilde{c}(0) \\ \frac{dq_2}{dt} &= -e q_2 - \mu q_2 - q_1 q_2 \tilde{c}(0) \\ \frac{d\tilde{g}_{11}}{dt} &= -2\mu \tilde{g}_{11} + 2e \tilde{g}_{12} - 2\tilde{g}_{11} q_2 \tilde{c}(0) - 2\tilde{g}_{12} q_1 \tilde{c}(0) \\ \frac{d\tilde{g}_{12}}{dt} &= (\tilde{g}_{11} - \tilde{g}_{12}) q_2 \tilde{c}(0) + (\tilde{g}_{22} - \tilde{g}_{12})(e - q_1 \tilde{c}(0)) - 2\mu \tilde{g}_{12} - q_1 q_2 \tilde{c}(0) \\ \frac{d\tilde{g}_{22}}{dt} &= -\frac{d\tilde{g}_{11}}{dt} - 2\frac{d\tilde{g}_{12}}{dt}, \end{aligned}$$

Setting  $dq_1/dt = 0 = dq_2/dt = d\tilde{g}_{11}/dt = \dots$  etc., and solving (which can be performed by Mathematica), we find the following equilibrium densities and correlation functions that will be needed later:

$$\begin{aligned} q_1^* &= (1 - P_0) Q_0 \\ q_2^* &= P_0 Q_0 \\ g_{12} + g_{22} &= \frac{Q_0 P_0 (1 - P_0)}{\phi(\omega)(1 + \rho) - (1 - P_0)} \end{aligned}$$

where

$$\begin{aligned} P_0 &= 1 - \frac{\mu(e + \mu)}{\tilde{c}(0)r} \\ Q_0 &= \frac{r}{\mu} \\ \phi(\omega) &= \frac{\tilde{c}(0)}{\tilde{c}(\omega)} \\ \rho &= \frac{1 - P_0}{1 + \frac{\epsilon}{\mu}} \end{aligned}$$

Derivation of connectivity–occupancy correlation: We are interested in the product-moment correlation  $R$  between occupancy  $\sigma(x)$  ( $= 1$  if there is an occupied patch at  $x$  and 0 if there is not) and connectivity  $S(x) = \sum_{y \in \gamma_A} h(x - y)$ , where  $\gamma_A$  is the set of all patches in the landscape whether occupied or not. This



can be defined as

$$R = \frac{\left\langle \left( \sigma(x) - \langle \sigma(x) \rangle_{x \in Y_A} \right) \left( S(x) - \langle S(x) \rangle_{x \in Y_A} \right) \right\rangle_{x \in Y_A}}{\left[ \left\langle \left( \sigma(x) - \langle \sigma(x) \rangle_{x \in Y_A} \right)^2 \right\rangle_{x \in Y_A} \left\langle \left( S(x) - \langle S(x) \rangle_{x \in Y_A} \right)^2 \right\rangle_{x \in Y_A} \right]^{1/2}},$$

where  $\langle \cdot \rangle_{x \in Y_A}$  denotes an average over all patches in the landscape, i.e.,

$\langle Z(x) \rangle_{x \in Y_A} = \frac{\sum_{x \in Y_A} Z(x)}{|Y_A|}$ . These averages were expressed as spatial moments (Supplementary Note 3 section 3.4), which were expanded using the perturbation expansion to leading order in  $\epsilon^d$  were retained. Expressions for  $q$  and  $g$  above were then substituted into the expression for  $R$  to give

$$R = \frac{\int \tilde{h}(\omega) (g_{12}(\omega) + g_{22}(\omega)) d\omega}{\sqrt{Q_0^3 P_0 (1 - P_0) \int \tilde{h}^2(\omega) d\omega}} \quad (7)$$

$$= \frac{\int_{\mathbb{R}^d} \frac{\tilde{h}(\omega) Q_0 P_0 (1 - P_0)}{\phi(\omega)(1 + \rho) - (1 - P_0)} d\omega}{\left( Q_0^3 P_0 (1 - P_0) \int_{\mathbb{R}^d} \tilde{h}^2(\omega) d\omega \right)^{\frac{1}{2}}} \quad (8)$$

Proof that landscape turnover weakens connectivity–occupancy correlation: From Eq. (7), we compute  $\partial R / \partial \rho$ , keeping the mean patch occupancy  $P_0$  and mean patch density  $Q_0$  constant:

$$\frac{\partial R}{\partial \rho} = - \frac{1}{\left( Q_0^3 P_0 (1 - P_0) \int_{\mathbb{R}^d} \tilde{h}^2(\omega) d\omega \right)^{\frac{1}{2}}} \int_{\mathbb{R}^d} \tilde{h}(\omega) \frac{\phi(\omega) Q_0 P_0 (1 - P_0)}{(\phi(\omega)(1 + \rho) - (1 - P_0))^2} d\omega.$$

This expression is always negative.

Kernel that maximises correlation: We take a variational approach and write  $\tilde{h}(\omega) = \tilde{h}_*(\omega) + \nu \delta(|\omega| - |\omega_1|)$ , where  $\delta$  is a Dirac Delta function. If the correlation is maximal when  $h = h_*$ , then  $\partial R / \partial \nu = 0$  for all  $\omega_1$ , which leads to

$$\tilde{h}_*(\omega) = \frac{1}{\phi(\omega)(1 + \rho) - (1 - P_0)}.$$

Variance and standard deviation of the optimal kernel: For any rotationally symmetric function  $M(x)$  with variance  $V_M = \int_{\mathbb{R}^d} x^2 M(x) dx / \int_{\mathbb{R}^d} M(x) dx$ , the Taylor expansion of the Fourier transform of  $M$  is

$$\tilde{M}(\omega) = \tilde{M}(0) (1 - a V_M \omega^2 + o(\omega^2)),$$

where  $a$  is a number that depends on spatial dimension, but not on  $M$ . Thus,  $\tilde{c}(\omega) = \tilde{c}(0) (1 - a V_c \omega^2 + o(\omega^2))$ , and expanding  $\tilde{h}_*$  gives

$$\tilde{h}_*(\omega) = \frac{A}{\rho + P_0} \left\{ 1 - a V_c \omega^2 \frac{1 + \rho}{\rho + P_0} + o(\omega^2) \right\}.$$

Therefore, the variance of the optimal connectivity kernel  $h_*$  is

$$\begin{aligned} V_* &= V_c \left( \frac{1 + \rho}{\rho + P_0} \right) \\ &= V_c \left( \frac{\tilde{c}(0)r + \mu^2}{\tilde{c}(0)r - e\mu} \right) \end{aligned}$$

Thus, the standard deviation of the optimal kernel is

$\lambda_* = l_c \sqrt{(\tilde{c}(0)r + \mu^2) / (\tilde{c}(0)r - e\mu)}$ , where  $l_c$  is the standard deviation of the colonisation kernel.

Parameters used in Fig. 3: **a**  $r = \mu = 0$ ,  $c(x) = 4 \cdot \text{Tophat}(x, 1)$  where Tophat  $(x, R)$  is a normalised tophat kernel with radius  $R$ ,  $e = 1$ . **b** Connectivity kernel  $h(x) = \exp(-x/\lambda)$ , colonisation kernel  $c(x) = 4 \cdot \text{Tophat}(x, l_c \sqrt{12})$  so that the colonisation and connectivity kernels have the same standard deviation when  $\lambda = l_c$ . Static landscape (solid lines, circles):  $\mu = r = 0$ ,  $e = 2$ ; dynamic landscape (dashed lines, triangles)  $\mu = r = 2$ ,  $e = 0$ ; **c**  $r = \mu$  and  $e = 2 - \mu$  so that patch density = 1 and mean-field occupancy  $P_0 = 1 - \frac{e(\mu + e)}{\tilde{c}(0)}$  is kept constant as  $\mu/e$  changes;  $c(x) = 4 \cdot \text{Tophat}(x, 4)$ ;  $h = \exp(-x/4)$ .

**Genetic similarity model.** Full operator  $L$  defining the model: Denoting two habitat types as species  $s_1$  and  $s_2$ , and denoting species with four different genotypes as  $s_3, s_4, s_5$  and  $s_6$ , the graphical definition of the model presented in Fig. 4a leads to the complete definition of the model by the following full operator  $L$ :

$$\begin{aligned} L &= \sum_{i=s_1, s_2} (L_i^{\text{IM}}(\kappa) + L_i^{\text{D}}(\mu)) + \sum_{j=s_3, s_4, s_5, s_6} \left[ L_j^{\text{D}}(m) + \sum_{k=s_3, s_4, s_5, s_6} L_{jk}^{\text{DE}}(c) \right] \\ &+ \sum_{i=s_3, s_4, s_5, s_6} \left[ \sum_{k_1=s_3, s_4, s_5, s_6} \left( \sum_{k_2=s_1, s_2} \phi_{k_1, k_2}^T \psi_{k_1, i}^{\text{BTF}} I_{ik_1, k_2}^{\text{BTF}}(r, d) \right) \right], \end{aligned}$$

where abbreviations have the following meaning: IM stands for Immigration, D for density-independent death, DE for death by external factor, BTF for birth to another type by facilitation (these processes are defined in Supplementary Note 1 section 1.3).

Probability densities: Using notations introduced in the main text, the probability density for any two individuals at locations  $y$  and  $x + y$  is considered as  $\sum_{i, i'} k_{i, i'}^{(2)}(y, x + y)$ , where the sum is taken over all possible pairs  $(i, i')$  of genotypes (there are four possible genotypes:  $\{A1, A2, B1, B2\}$ ). Probability density for individuals at  $y$  and  $x + y$  with allele  $j$  in neutral locus (i.e.,  $j = 1, 2$ ) is defined as  $\left( k_{A1, A1}^{(2)}(\bullet) + k_{A1, B1}^{(2)}(\bullet) + k_{B1, A1}^{(2)}(\bullet) + k_{B1, B1}^{(2)}(\bullet) \right)$ , and with allele  $\beta$  in selective locus (i.e.,  $\beta = A, B$ ) is defined as  $\left( k_{\beta 1, \beta 1}^{(2)}(\bullet) + k_{\beta 1, \beta 2}^{(2)}(\bullet) + k_{\beta 2, \beta 1}^{(2)}(\bullet) + k_{\beta 2, \beta 2}^{(2)}(\bullet) \right)$ , where in both cases  $\bullet$  denotes locations of individuals  $\bullet = \{y, x + y\}$ .

Analytical expressions of similarity functions: We use Model Constructor to obtain expressions for  $q$ ,  $p$  and  $g$  for the model, and hence expressions for the leading terms in the expansion for the two-point correlation functions (second spatial moments)  $k_{ij}^{(2)}$ . To first-order in  $\epsilon^d$ , and from now on assuming we are in spatial dimension  $d = 2$ , the similarity functions  $F_n(x)$  and  $F_s(x)$  take the form

$$\begin{aligned} F_n(x) &= \frac{1}{2} + \frac{\pi}{4q^*} \int_0^\infty \frac{(1 - 2\nu)\tilde{d}(\omega)}{\tilde{d}(0) - (1 - 2\nu)\tilde{d}(\omega)} J_0(2\pi x\omega) \omega d\omega, \\ F_s(x) &= F_n(x) + \tau^2 \int_0^\infty \frac{(1 - 2\nu)\tilde{d}(\omega)}{\tilde{d}(0) - (1 - 2\nu)\tilde{d}(\omega)} \frac{\pi(1 - 2\nu)f_0\tilde{d}(\omega)^2(\omega)/\tilde{r}(0)}{2f_0\kappa[\tilde{d}(0) - (1 - 2\nu)\tilde{d}(\omega)]\tilde{r}(0) + \mu^2} \\ &\quad \times J_0(2\pi x\omega) \omega d\omega, \end{aligned}$$

where  $J_0(\omega)$  is Bessel function of the first kind of order zero and

$$q^* = \frac{2f_0\kappa\tilde{d}(0)\tilde{r}(0) - m\mu}{4\mu\tilde{c}(0)}$$

is the mean-field density for any single genotype.

In the limit of very rare mutations:  $\nu \rightarrow 0$ , the denominator in the integrals above is dominated by the behaviour for  $\omega$  small. In small- $\omega$  limit in 2D one obtains  $\tilde{d}(\omega) \approx \tilde{d}(0)(1 - 2\pi^2\sigma^2\omega^2) + o(\omega^3)$ , where  $\sigma^2$  is the variance of the kernel  $d$ . As a result, we have

$$\begin{aligned} F_n(x) &\approx \frac{1}{2} + \frac{1}{8\pi q^* \sigma^2} \int_0^\infty \frac{J_0(2\pi x\omega) \omega d\omega}{R_\nu^2 + \omega^2} \\ &= \frac{1}{2} + \frac{1}{8\pi q^* \sigma^2} K_0(2\pi x R_\nu) \\ &= \frac{1}{2} + \frac{1}{2N_e \pi \sigma^2} K_0(x/x_n), \quad x_n = \frac{\sigma}{2\sqrt{\nu}} \end{aligned}$$

where

$$R_\nu^2 = \frac{\nu}{\pi^2 \sigma^2},$$

$N_e = 4q^*$ , and we have assumed  $x \ll \sigma$ ;  $K_0$  is a modified Bessel function of the second kind of order zero. Similarly, when  $\nu \rightarrow 0$ , the similarity at the selective locus is

$$F_s(x) \approx F_n(x) + \frac{\tau^2}{8\nu q_h \pi \sigma^2} \frac{K_0(x/x_n) - K_0(x/x_s)}{\mu/M}, \quad x_s = \frac{x_n}{\sqrt{1 + \mu/M}},$$

where  $q_h = \kappa/\mu$  is the density of habitat patches of a single type, and  $M = 4\nu q_h f_0 \tilde{r}(0) \tilde{d}(0)$  is the rate with which an existing individual produces a new mutant.

Derivation of length scales: We define the length scale for a similarity function as the mean distance over which it decays to its asymptotic value at large distances ( $\lim_{x \rightarrow \infty} F(x) = \frac{1}{2}$ ):

$$l = \frac{\int_0^\infty x(F(x) - \frac{1}{2}) dx}{\int_0^\infty (F(x) - \frac{1}{2}) dx}.$$

According to this definition, using the small- $\nu$  expressions for  $F_n$  and  $F_s$  in terms of Bessel functions, we obtain the expression for the ratio of length scales presented in the main text.

Parameters used in Fig. 4: **b**  $\nu = 0.01$ ,  $\kappa = 0.05$ ,  $\mu = 1$ ,  $\tau = 1$ ,  $f_0 = 1$ ,  $m = 1$ ;  $r(x) = 3.5 \times \text{Tophat}(x, 2)$ ;  $d(x) = 3 \times \text{Tophat}(x, 2)$ ;  $c(x) = 1.25 \times \text{Tophat}(x, 3/20)$ , where Tophat  $(x, R)$  is a normalised tophat kernel with radius  $R$ . **c**  $d(x) = r(x) = \text{Tophat}(x, 6)$ , and  $c(x) = \text{Tophat}(x, 20)$ .

**Optimal foraging model.** The model is specified in section ‘Optimal foraging’ in the paper and Fig. 5a, and is defined by the operator

$$\begin{aligned} L &= L_1^{\text{IM}}(b) + L_1^{\text{D}}(h) + L_{21}^{\text{BT}}(h\lambda r) + L_2^{\text{D}}(\mu) \\ &+ L_3^{\text{I}}(m_\text{S}c) + L_4^{\text{I}}(m_\text{FC}c) + L_{23}^{\text{CT}}(\alpha) + L_{23}^{\text{DE}}(\gamma f) + L_{34}^{\text{CTC}}(\gamma f), \end{aligned}$$

where the entity types are defined as follows: 1: resource target generators; 2: resource targets; 3: slow-moving consumer; 4: fast-moving consumer; the abbreviation BT stands for birth to another type, J for jump, CTC for change in type by consumption, other abbreviations were already explained above (the full definition is presented in Supplementary Note 1 section 1.3, see also Supplementary Tables 1–3). Parameter  $h$  determines the rate at which target generators disappear; we will study the case where clusters of targets are created simultaneously by taking the



limit  $h \rightarrow \infty$ . Since each target generator creates resources at rate  $h\lambda$ , and has mean lifetime  $1/h$ , it generates a Poisson distributed number of resource targets with mean  $\lambda$ . This is a convenient method for generating clusters of individuals from an underlying process that generates individuals; a similar approach was used in ref. <sup>15</sup> to generate clusters of habitat patches.

Our interest is in the equilibrium rate  $\rho^*$  by which each searcher consumes targets. This can be computed by solving the stationary state of the system under the condition that the density of searchers, denoted here by  $A$ , is constant. This yields  $\rho^* = (\beta/k_2^* - \mu)k_2^*/A$ , where  $\beta = \lambda b$  is the appearance rate of targets,  $k_2^*$  is the stationary density of targets and the expression in brackets is the rate at which a randomly selected target disappears due to consumption. As we are not interested in resource competition among multiple searchers, we consider the case of a single searcher, obtained technically by taking the limit  $A \rightarrow 0$ . Note that, in the main text, we have set  $\epsilon = 1$ , i.e., the length scale of the kernels  $f$  and  $r$  is the true biological one and has not been rescaled by  $\epsilon$ .

**Resource consumption rate:** Using Model constructor, the stationary density of targets in the mean-field approximation is  $k_2^* \approx q_2^* = \beta/(A\gamma + \mu)$ , and thus the consumption rate of targets by a single searcher is  $\rho^* \approx \rho_0 = \gamma\beta/\mu$  independently of the parameter  $\alpha$ .

In the first-order approximation, the stationary density of targets is  $k_2^* \approx q_2^* + \epsilon^d p_2^*$ . The general expression for  $p_2^*$  is given using ‘The model constructor’ toolbox:  $p_2^* = \int_0^\infty \text{IntF}(\omega) d\omega$ , where the integrand function  $\text{IntF}(\omega)$  is

$$\text{IntF}(\omega) = -\frac{2\gamma\omega\tilde{f}(\omega)}{(A\gamma + \mu)} (\tilde{g}_{23}(\omega) + \tilde{g}_{24}(\omega)).$$

Using ‘The model constructor’ toolbox to calculate the limit  $h \rightarrow \infty$  (after which the results no longer depend on  $h$ ) and to extract the leading term in the limit  $A \rightarrow 0$ , which is achieved by the following Mathematica commands:

```
IntF = IntF /. h -> (1/ih);
IntF = Normal[Series[IntF, {A, 0, 1}]];
IntF = Limit[IntF, ih -> 0]
```

We thus obtain:

$$\text{IntF}(\omega) = \frac{2YAb\gamma^2\omega\tilde{f}(\omega)^2}{\mu^2(b\gamma\lambda + \alpha\mu)Z},$$

where

$$\begin{aligned} Y &= b^2\gamma^2\lambda^3 + b\gamma\lambda^2\mu(2\alpha + m_F + \mu) + \lambda\alpha\mu^2(\alpha + m_S + \mu) \\ &\quad + (-m_F + m_S)\tilde{r}(\omega)^2\mu^2\alpha(\alpha - \tilde{c}(\omega))^2 \\ &\quad - \mu\tilde{c}(\omega)\lambda(b\gamma\lambda m_F + \alpha m_S\mu); \\ Z &= b\gamma\lambda(m_S + \mu) + \mu(m_F + \mu)(\alpha + m_S + \mu) \\ &\quad - (b\gamma\lambda m_S + \mu(\alpha m_F + 2m_F m_S + m_F\mu + m_S\mu))\tilde{c}(\omega) \\ &\quad + m_F m_S \mu \tilde{c}(\omega)^2. \end{aligned}$$

The expressions for  $\tilde{g}_{ij}$  are still rather unwieldy, but we can simplify them by considering the limit where in the fast movement mode the movements are very fast ( $m_F \rightarrow \infty$ ), and the slow movements are very slow ( $m_S = 0$ ). This is achieved by the Mathematica command

```
Simplify[Limit[IntF, mf -> Infinity]] /. ms -> 0
```

resulting in

$$\text{IntF}(\omega) = \frac{2Ab\gamma^2\omega\tilde{f}(\omega)^2(b\gamma\lambda^2 - \alpha^2\mu\tilde{r}(\omega)^2)}{\mu^2(\alpha + \mu)(b\gamma\lambda + \alpha\mu)}.$$

At this limit, the first-order approximation for the consumption rate of targets is  $\rho^* \approx \rho_0 + \epsilon^d \rho_1$ , where

$$\rho_1 = \frac{\beta\gamma^2(\alpha\lambda\mu I_1 - \beta\gamma I_2)}{\mu(\beta\gamma + \alpha\mu)(\alpha + \mu)},$$

and  $I_1 = \int_0^\infty 2\pi\omega\tilde{f}(\omega)^2\tilde{r}(\omega)^2 d\omega$  and  $I_2 = \int_0^\infty 2\pi\omega\tilde{f}(\omega)^2 d\omega$ .

**Optimal switching rate:** The optimal switching rate is found by the value of  $\alpha$  where  $\partial\rho_1/\partial\alpha = 0$ , which from the above expression for  $\rho_1$  is

$$\alpha^* = \rho_0 \frac{I_2}{I_1} + \sqrt{\rho_0 \left(1 + \frac{\rho_0}{\mu} \frac{I_2}{I_1}\right) \left(1 + \frac{I_2}{\lambda I_1}\right)}.$$

**Threshold cost of fast foraging:** During the fast-foraging mode, the resource acquisition rate is  $\rho_0$ , so it takes a time  $1/\rho_0$  before the consumer finds a resource and switches to the slow mode. If the fast mode costs  $\kappa$  resource units per unit time, then the rate of resource expenditure due to fast foraging is  $\alpha\kappa/\rho_0$ . The net resource gain is

$$\rho_0 + \epsilon^d \rho_1 - \alpha\kappa/\rho_0 = \rho_0 + \epsilon^d \frac{\rho_0\gamma(\alpha\lambda I_1 - \rho_0 I_2)}{(\rho_0 + \alpha)(\alpha + \mu)} - \frac{\alpha\kappa}{\rho_0},$$

where we used  $\rho_0 = \gamma\beta/\mu$ . This function has a maximum at  $\alpha > 0$  (and,

correspondingly, its derivative with respect to  $\alpha$  at  $\alpha = 0$  is positive) unless

$$\kappa > \epsilon^d \frac{\rho_0\gamma}{\mu} \left[ \lambda \int_{\mathbb{R}^d} \{\tilde{f}(\omega)\tilde{r}(\omega)\}^2 d\omega + \left(1 + \frac{\rho_0}{\mu}\right) \int_{\mathbb{R}^d} \tilde{f}^2(\omega) d\omega \right],$$

in which case the fastest net resource acquisition is at  $\alpha = 0$ .

Parameters used in Fig. 5: Parameters not varied in the panels are set to  $\gamma = 1$ ,  $\beta = 0.01$ ,  $\mu = 0.1$ ,  $\lambda = 10$ ,  $f(x) = r(x) = c(x) = \text{Tophat}(x, 5)$ . For panel c, the theoretical curve is calculated for  $h = m_F = 1/m_S \rightarrow \infty$ , whereas the simulations used  $h = m_F = 1/m_S = 1000$ .

**Reporting summary.** Further information on research design is available in the Nature Research Reporting Summary linked to this article.

## Data availability

Simulation data used to generate Figs. 3, 4 and 5, created using the Model Simulator toolkit, are available from Figshare at <https://doi.org/10.6084/m9.figshare.9632531>.

## Code availability

Model Constructor and Model Simulator are released under the GNU Public License v. 2 and are available from Figshare at <https://doi.org/10.6084/m9.figshare.9633161>. Included with the software are sample files and code for running the case studies in this paper.

Received: 26 September 2017; Accepted: 22 August 2019;

Published online: 17 October 2019

## References

- Grimm, V. & Railsback, S. F. *Individual-based Modeling and Ecology*. (Princeton University Press, 2005).
- Filatova, T., Verburg, P. H., Parker, D. C. & Stannard, C. A. Spatial agent-based models for socio-ecological systems: challenges and prospects. *Environ. Model. Softw.* **45**, 1–7 (2013).
- DeAngelis, D. L. & Yurek, S. Spatially explicit modeling in ecology: a review. *Ecosystems* **20**, 1–17 (2017).
- Wallentin, G. Spatial simulation: a spatial perspective on individual-based ecology—a review. *Ecol. Modell.* **350**, 30–41 (2017).
- MacArthur, R. & Levins, R. Competition, habitat selection, and character displacement in patchy environments. *Proc. Natl Acad. Sci. USA* **51**, 1207–1210 (1964).
- May, R. M. & Anderson, R. M. Regulation and stability of host-parasite population interactions II. Destabilising processes. *J. Anim. Ecol.* **47**, 249–267 (1978).
- Durrett, R. & Levin, S. A. The importance of being discrete (and spatial). *Theor. Popul. Biol.* **46**, 363–394 (1994).
- Neuhauser, C. Mathematical challenges in spatial ecology. *Not. AMS* **48**, 1304–1314 (2002).
- Mahmutovic, A., Fange, D., Berg, O. G. & Elf, J. Lost in presumption: stochastic reactions in spatial models. *Nat. Methods* **9**, 1163–1166 (2012).
- Macal, C. M. Everything you need to know about agent-based modelling and simulation. *J. Simul.* **10**, 144–156 (2016).
- Ovaskainen, O. et al. A general mathematical framework for the analysis of spatiotemporal point processes. *Theor. Ecol.* **7**, 101–113 (2014).
- Bolker, B., Pacala, S. W. & Levin, S. A. Moment methods for ecological processes in continuous space. In *The Geometry of Ecological Interactions: Simplifying Spatial Complexity* (eds Dieckmann, U., Law, R. & Metz, J. A. J.) 388–411 (Cambridge University Press, 2000).
- Ovaskainen, O. & Cornell, S. J. Space and stochasticity in population dynamics. *Proc. Natl Acad. Sci. USA* **103**, 12781–12786 (2006).
- Plank, M. J. & Law, R. Spatial point processes and moment dynamics in the life sciences: a parsimonious derivation and some extensions. *Bull. Math. Biol.* **77**, 586–613 (2015).
- Cornell, S. J. & Ovaskainen, O. Exact asymptotic analysis for metapopulation dynamics on correlated dynamic landscapes. *Theor. Popul. Biol.* **74**, 209–225 (2008).
- Moilanen, A. & Nieminen, N. Simple connectivity measures in spatial ecology. *Ecology* **83**, 1131–1145 (2002).
- Kool, J. T., Moilanen, A. & Treml, E. A. Population connectivity: recent advances and new perspectives. *Landscape Ecol.* **28**, 165–185 (2013).
- Hodgson, J. A., Moilanen, A. & Thomas, C. D. Metapopulation responses to patch connectivity and quality are masked by successional habitat dynamics. *Ecology* **90**, 1608–1619 (2009).

19. Hodgson, J. A., Thomas, C. D., Wintle, B. A. & Moilanen, A. Climate change, connectivity and conservation decision making: back to basics. *J. Appl. Ecol.* **46**, 964–969 (2009).
20. Felsenstein, J. A. Pain in the torus: some difficulties with models of isolation by distance. *Am. Nat.* **109**, 359–368 (1975).
21. Kimura, M. & Weiss, G. H. The stepping stone model of population structure and the decrease of genetic correlation with distance. *Genetics* **49**, 561–576 (1964).
22. Chave, J. & Leigh, G. E. A spatially explicit neutral model of beta-diversity in tropical forests. *Theor. Popul. Biol.* **62**, 153–168 (2002).
23. Barton, N. H., Kelleher, J. & Etheridge, A. M. A new model for extinction and recolonization in two dimensions: quantifying phylogeography. *Evolution* **64**, 2701–2715 (2010).
24. Barton, N. H., Depaulis, F. & Etheridge, A. M. Neutral evolution in spatially continuous populations. *Theor. Popul. Biol.* **61**, 31–48 (2002).
25. Etheridge, A. M. Survival and extinction in a locally regulated population. *Ann. Appl. Probab.* **14**, 188–214 (2004).
26. Viswanathan, G. M. et al. Optimizing the success of random searches. *Nature* **401**, 911–914 (1999).
27. Sims, D. W. et al. Scaling laws of marine predator search behaviour. *Nature* **451**, 1098–1102 (2008).
28. Fryxell, J. M. et al. Multiple movement modes by large herbivores at multiple spatiotemporal scales. *Proc. Natl Acad. Sci. USA* **105**, 19114–19119 (2008).
29. Pyke, G. H. Understanding movements of organisms: It's time to abandon the Lioevy foraging hypothesis. *Methods Ecol. Evol.* **6**, 1–16 (2015).
30. Drawert, B. et al. Stochastic simulation service: bridging the gap between the computational expert and the biologist. *PLOS Comput. Biol.* **12**, e1005220 (2016).
31. Finkelshtein, D. L. On convolutions on configuration spaces. I. Spaces of finite configurations. *Ukr. Math. J.* **64**, 1752–1775 (2013).
32. Wolfram Research, Inc., Mathematica, Version 10.4 Champaign, IL (2018).
33. Gillespie, D. Exact stochastic simulation of coupled chemical reaction. *J. Phys. Chem.* **81**, 2340–236 (1977).

## Acknowledgements

O.O. and P.S. were supported by funding from the Academy of Finland (Grants nos. 1273253 and 250444 to O.O.) and the Research Council of Norway (CoE grant no. 223257). We would like to thank Mike Begon and Jane Rees for comments on the paper.

## Author contributions

S.C., O.O. and Y.S. designed the study; S.C., Y.S. and O.O. wrote the main text of the paper. Case study 1 was performed by S.C., case study 2 by Y.S. and S.C., case study 3 by

O.O., S.C. and Y.S. In Supplementary Notes: Y.S. and D.F. wrote Supplementary Note 1 sections 1.1–1.3 (Introduction to mathematical framework and derivations for selected basic processes); Y.S. wrote Supplementary Note 1 sections 1.4 and 1.5 (derivations for a general process), Supplementary Note 2 sections 2.1 and 2.3 (Mathematica code and its tutorial); P.S. wrote Supplementary Note 2 section 2.2 (Simulation code and its tutorial); S.C. and Y.S. wrote Supplementary Note 3; Y.S. and S.C. wrote Supplementary Note 4; O. O. and Y.S. wrote Supplementary Note 5. All authors contributed to discussions and edited the paper.

## Competing interests

The authors declare no conflict of interest.

## Additional information

**Supplementary information** is available for this paper at <https://doi.org/10.1038/s41467-019-12172-y>.

**Correspondence** and requests for materials should be addressed to S.J.C.

**Peer review information** *Nature Communications* thanks the anonymous reviewers for their contribution to the peer review of this work. Peer reviewer reports are available.

**Reprints and permission information** is available at <http://www.nature.com/reprints>

**Publisher's note** Springer Nature remains neutral with regard to jurisdictional claims in published maps and institutional affiliations.



**Open Access** This article is licensed under a Creative Commons Attribution 4.0 International License, which permits use, sharing, adaptation, distribution and reproduction in any medium or format, as long as you give appropriate credit to the original author(s) and the source, provide a link to the Creative Commons license, and indicate if changes were made. The images or other third party material in this article are included in the article's Creative Commons license, unless indicated otherwise in a credit line to the material. If material is not included in the article's Creative Commons license and your intended use is not permitted by statutory regulation or exceeds the permitted use, you will need to obtain permission directly from the copyright holder. To view a copy of this license, visit <http://creativecommons.org/licenses/by/4.0/>.

© The Author(s) 2019